

I. Introduction

1. The tasks in sentence processing:
 - recognising words
 - determining syntactic relationships between them
 - determining semantic relationships between them
 - interpreting contextually-dependent information
2. There is lots of evidence that all of these are performed on-line
 - words are recognised as soon as hearer has enough phonetic input to tell them apart from probable alternatives
 - lexical, syntactic, contextual information plays a role
 - so, recognising words is not done as an isolated/modular task
 - so, hearers don't wait for the end of the word
 - syntactic relationships are decided at the first moment possible
 - evidence of decisions in syntactic ambiguity appears on first fixation upon disambiguating word
 - processing effects in anomalous sentences immediately or soon after the word at which anomaly occurs
 - interpreting referential expressions from the context begins immediately
 - can influence syntactic ambiguity resolution
3. Methods used to study processing reflect its on-line nature
 - monitoring
 - probing and priming
 - self-paced reading
 - eye-tracking
 - event-related potentials
4. On-line syntactic processing – the issues
 - lots of temporary ambiguities
 - words are frequently ambiguous as to syntactic category: *Time flies!*
 - Syntactic structure is hierarchical & recursive, so many relationships between non-adjacent words
 - syntactic ambiguities = different semantic & discourse results
 - semantics and discourse can provide constraints to resolve ambiguities
 - e.g., thematic information: good & bad Agents
 - e.g., how new events fit into discourse
 - so, how are these constraints used?
 - Theory 1: rapid, optimal use of different sources of information: correlated constraints provide evidence for/against competing alternatives
 - Theory 2: isolated modules process information from different sources separately (at least initially), allowing for faster processing
5. Summary of issues in studying sentence comprehension:
 - modularity debate
 - what is the nature of linguistic representations involved?
 - how this grammatical knowledge is accessed & used in processing?

II. Historical overview

1. In 1960s: George Miller & students – Chomskian derivation-based approach

- “language as product” = getting linguistic representations – main goal of processing
- Chomsky's theories – phrase structure rules build *deep structure* that reflects meaning, transformations change these into the pronounced *surface structure*
- Miller's main hypothesis – hearer builds surface structure representations, then reconstructs the deep structure by “undoing” transformations
- Difficulty of processing affected by several factors
 - number of transformations
 - limited working memory – how much must be remembered during processing
- Experimental evidence – consistent with deep-structure as product to be remembered & also consistent with propositional content as product
 - Sachs 67 – people remembered gist, not structure
 - Blumenthal 67 & Wanner 74/77 – effectiveness of a word for helping to remember a sentence ~ number of times the word appears in deep structure/proposition
- 2. In the 70s – transition from DTC to modern research
 - deep structure not retained, but rather content that's relevant to context
 - difficulty depends also on prior context (Greene 1970) & semantics of NPs (Slobin 1966)
 - Perceptual theories (esp. Bever 1970) – how surface is translated into deep structure without transformations
 - e.g. NV(N) -> SubjectVerb(Object)

initial misparsing = “garden path” sentences

The horse raced past the barn fell

- made use of various sources of information
- local ambiguities become central to study of processing
- clauses as units of processing
 - clicks within words remembered as being at clause boundaries
 - reaction times to a tone slower at the end of a clause, faster at the beginning
 - multiple readings more available within clause than across clauses
 - memory for surface form declined after clause boundary
- while clauses are important, important processing also happens prior to clause boundaries
 - 1980s & 1990s – renewed optimism that grammatical representations participate in processing
 - Bresnan 1978 – strong competence hypothesis
 - More syntax in lexical representations (argument structures)
 - syntactic structure not a product of rules but unordered local constraints

III. Ambiguity resolution and current issues in sentence processing

1. Syntactic category ambiguity
 - senses of words from both categories initially activated – multiple access
 - consistent with lexical access as separate module
 - or because syntax is only a weak constraint - under strong syntactic constraints, only the biased sense primes (*would* vs *wood* in “would” contexts)
 - Relative frequency of alternatives often changes dramatically with the preceding context
 - Juliano & Tanenhaus 1993 – effects of context on initial preferences for *that* – biases depended on frequency-in-a-particular-syntactic environment, not simple lexical frequency
 - Timing of resolution

- Frazier & Rayner 1987 – *The/These warehouse fires...* parser delays interpreting
 - ambiguous regions have faster fixation during NP
 - slower fixation during the rest of the sentence
- BUT – MacDonals 1993 – same pattern arises without ambiguities b/c use of *these* requires special referential context (cross-modal naming & self-paced reading)
- Global ambiguities – *Visiting relatives can be boring*

2. Attachment ambiguity

- Three related questions
 1. what alternatives are available at the point of the ambiguity?
 2. what degree of commitment is made to one or more alternatives?
 3. What information is used to guide the commitments?
- Possible answers
 1. delaying commitments until enough information accrues to make an informed decision (e.g. Marcus 1980)
 2. making partial syntactic commitments that are consistent with upcoming alternatives (Marcus, Hindle, & Fleck 1983; Weinberg in press)
 3. making an initial commitment to a single analysis using only syntactic category information – *garden-path model* (Frazier & Fodor 1978, Kimball 1973)
- choices made on syntactic basis like “avoid adding new nodes”
- Frazier & Fodor framework
 - principle of Minimal Attachment (build the simplest structure)
 - principle of Late Closure – recent attachment preferred when competing structures are equally simple
- Modules – each computes a different type of representation
 - constituent-building module – reduced memory load
 - if initial structure incorrect, it is revised creating temporary increase in processing difficulty
 - revisions guided by “thematic” & discourse information
- 4. making a single commitment guided by lexical representations (Ford, Bresnan & Kaplan 1982)
- 5. developing multiple analyses in parallel (Gorrell 1989, Kurtzman 1985) with selection based on
 - trying to satisfy grammatical principles (Pritchett 1992),
 - pragmatic fit to the discourse model – *referential theory* (Altmann & Steedman 1988, Crain & Steedman 1985)
 - principle of Minimal Presupposition w.r.t. discourse model – explains initial preferences without context
 - presuppositional complexity – finding or creating a referent for definite NPs & other referring expressions
 - choices are guided by best fit to pragmatic constraints
 - compatible with modules, labeled “weakly interactive”
 - multiple constraints including memory – *constraint-based models* (Bates & MacWhinney 1989, Gibson 1991, Marslen-Wilson 1973, Oden 1978, Taraban & McClelland 1988)
 - competition among incompatible alternatives, guided by satisfaction of constraints
 - processing difficulty when information is inconsistent & when input is

- inconsistent with previously biased alternative
 - no sharp break between building and evaluating structure
- generally underspecified, so hard to make predictions
- this is being addressed in *constraint-based lexicalist* models (e.g. McDonald et al. 1994)
 - attachment ambiguities arise from intersection of several types of ambiguities
 - argument structure ambiguities
 - syntactic category ambiguities
 - morphological ambiguities
 - access of lexical representations is frequency weighted
 - contextual effects mediated by lexical representations

Which charity did John donate...

Horse raced past the barn fell vs Land mine hidden in the sand exploded

Empirical studies

- structural preferences
 - Minimal Attachment & Late Closure supported by series of eye-tracking experiments (Frazier & Rayner 1982 and subsequent studies, Ferreira & Clifton 1986, Trueswell et al 1994) – readers experience processing difficulty as soon as disambiguating evidence is encountered favouring nonminimal attachment
 - *The pupil spotted by the proctor was expelled*
 - *The pupil who was spotted by the proctor was expelled*
 - *The pupil seen by the proctor was expelled*
 - Fixation durations increase in PP
 - higher likelihood of eye regression
 - Ambiguous vs unambiguous showed no systematic differences in overall processing – support for initial single analysis (if building parallel structures is costly) – very different from lexical ambiguity resolution
 - Initial attachments may be unaffected by lexical-syntactic information, local semantic context, & discourse context
 - Important exceptions to Late Closure in Spanish (Cuetos & Mitchell 1988) – frequency sensitive attachment preferences? Recency vs preference weighted differently? Arguments & non-arguments processed differently?
- Verb information
 - when is verb-specific information available? - both subcategorisation and thematic information available as soon as a verb is encountered
 - multiple argument structures increase processing complexity – lexical decision times in cross-modal task (Shapiro Zurif & Grimshaw 1989)
 - processing more complex when less preferred argument structure used (clifton, Frazier & Connine 1984)
 - McElree 1993, McElree & Griffith 1995 – category & subcategory information activated equally quickly, with less frequent subcategorisation info as rapid but maybe not as strong as more frequent ones – lexically based syntactic alternatives available in parallel
 - when verb preferences & Minimal attachment conflict
 - *After the child visited the doctor prescribed a course of injections*
 - *The student forgot the solution was in the back of the book*
 - *After the child sneezed the doctor prescribed a course of injections*
 - *The student hoped the solution was in the back of the book*
 - lexical filtering of syntactic alternatives – category-based NP attachment

- followed by rapid lexical filtering
 - constraint-lexicalist theories – category-based generalisation + most frequent attachment in the set of words
 - recent eye-tracking studies & connectionist simulations confirmed that more lexical item-specific frequencies play a role – so support for constraint-based lexicalist view
- Thematic fit - Inanimate nouns are bad Agents
 - Ferreira and Clifton 1986 – varied animacy of subject NP in reduced & unreduced relative clauses
 - first- and second-pass reading times unaffected by animacy, so support independence of initial parsing decisions
 - fixation times at the verb longer for inanimates – so aware of the implausibility of inanimate Agent
 - Trueswell et al 1994 – clear effects of animacy on PP-attachment disambiguation
 - differences between weakly & strongly constraining nouns
 - for strongly constraining nouns – no difficulty for reduced relatives compared to unambiguous controls
 - Thematic constraints interact with argument structure & postambiguity constraints, also with availability of the syntactic alternatives, show frequency effects (stronger thematic effects for verbs that are often used in passive participle, weakest for those typically used in simple past tense)
- Referential context
 - Some studies find immediate effects of referential context reducing or eliminating difficulty for less preferred structure
 - Other studies found only weak or delayed effects of referential context
 - Generalisations -
 - effects found most consistently with PP attachment – alternatives have most similar frequencies
 - weaker effects in reduced relative clauses where the frequencies are different
 - interaction with availability of of dispreferred structure (how strongly verb requires locative argument in *put away the book on the top shelf/the Civil War*) & with local semantic constraints (weaker effect for action verbs with strong VP attachment preference & stronger for psych/perception verbs with slight NP preference, as in *the thief blew up the/a safe with...*)
 - individual differences – high vs low memor-span for keeping parallel possibilities in mind
 - So – referential context can provide constraints that are rapidly used. But parsing preferences cannot be reduced to discourse factors
 - Research showed that readers incrementally update their discourse models to relate referential expressions to entities in their model, including events
- Intonation & prosody
 - duration of a segment increases at major syntactic boundary; pitch tends to decrease at major syntactic boundary, so resolves ambiguities in boundary placement
 - prosody can override late closure preferences
 - prosody can override verb-subcategorisation preferences
 - prosody provides evidence for referential context – contrastive focus etc.

--- enough for now; we will revisit other parts of this paper later in the semester

