It’s important to consider the role of working memory in sentence processing:

- If language is **discontinuous**, expect it to be processed idiosyncratically. Domain-specific: less preferable b/c requires more cognitive machinery.
- If language is **continuous**, expect it to be processed like other stimuli. Domain-general: requires less cognitive machinery, a priori preferable.

How do we process other stimuli?

Consider language to be specialization of hierarchic sequential prediction. We can remember a few things over short periods if they are not connected:

- You notice Mom making breakfast.
- As part of that, she’s preparing coffee.
- Then there are no filters, so she’s looking for them in the pantry.
- Then the phone rings, and she starts talking to Ms. Friendly.
- Now she hangs up, asks ‘what was I doing?’ (You have to remember!)

Maybe sent. processing is similar: ‘if .. either .. the car dies ..’ (or?/then?)
This kind of memory had been modeled as a fixed-capacity buffer:

- George Miller ’56: \(7 +/– 2\) elements
- Nelson Cowan ’01: \(4 +/– 1\) element

Like a computer: if you use all your elements, you have to forget something.

These were called **dual store** models:

1. long-term store: cued access – think of Mom and get name, related info
2. short-term store: serial access – last few ideas in order, even if unrelated

Problem with this story: we can remember lots of unrelated sequences:

- people that walked past you during the last few minutes
- what your lectures have covered over the last few hours
- what TV shows you watched over the last few days
- what courses you have taken over the last few semesters

Multi-scale sequences coexist, don’t displace each other (Robert Crowder ’82).
The Role of Working Memory in Sentence Processing

Evidence like this instead favors a **unitary store**:

- cued by **content-based features** for things like mother’s name,
- cued by **temporal features** for things like people walking past.

Both can be short-/long-term (but content-based associations accumulate; whereas noisy changes in temp feats makes past events harder to distinguish).

Content-based/temporal cueing can happen in sentence processing as well:

- content-based cueing tells you after ‘the’ to expect a noun,
- temporal cueing tells you what speaker was saying before noun phrase.

This explanation **may** predict a cost to **encode** or **integrate** referents:

- \[\text{encode} \begin{cases} \text{v} \text{ if (encode) } v \text{ either (encode) } v \text{ the car dies } \text{ (integrate) or } \cdots \end{cases} \]

and **may** predict cost to **maintain** stored referents (e.g. at ‘car’) — decay; but **does not** predict a cost for **alternative** hypotheses — no temporal cue! (consistent with ‘leopard/poacher’ findings by van Gompel & al ’01)
The Role of Working Memory in Sentence Processing

Evidence for integration/maintenance cost relies on filler-gap dependencies...

Ted Gibson '98: evidence for integration cost

- **stimuli**: sentences in self-paced reading
  - (a) ‘The reporter who attacked the senator admitted the error.’ (nested)
  - (b) ‘The reporter who attacked the senator admitted the error.’ (not nested)

- **measure**: self-paced reading durations

- **results**: large delay at ‘admitted’ at end of nesting (a)

Ted Gibson '98: evidence for maintenance cost

- **stimuli**: sentences in self-paced reading
  - (a) ‘The witness who the evidence implicated seemed nervous.’ (nest)
  - (b) ‘The witness thought the evidence implicated his neighbor.’ (no nest)

- **measure**: self-paced reading durations

- **results**: large delay in ‘...’ (‘examined by the lawyer’) within nesting (a)
Unbounded Dependencies (‘Filler-gap Constructions’)

What are filler-gap dependencies?

Argument ‘-g’ that can pass down through final children: \( \phi^{-gN} \rightarrow \psi \chi^{-gN} \)
satisfying argument ahead/behind predicate: \( \phi^{-gN} \rightarrow \phi^{-aN} / \phi^{-gN} \rightarrow \phi^{-bN} \)
What are filler-gap dependencies?

Argument ‘-g’ can also pass down through initial children: \( \phi-gN \rightarrow \psi-gN \chi \)
satisfying argument ahead/behind predicate: \( \phi-gN \rightarrow \phi-aN / \phi-gN \rightarrow \phi-bN \)
Gaps also occur in relative clauses (with relative pronoun arguments ‘-r’)
(rules to introduce relative pronoun: \( N \rightarrow N \, C-rN, C-rN \rightarrow N-rN \, V-gN \)):
Relativizers / zero relatives: $\text{C-rN} \rightarrow \text{C-bV V-gN} / \text{C-rN} \rightarrow \text{N V-aN-gN}$
Gaps also show up in content questions (with interrogative ‘-i’ arguments) (this also requires subject-auxiliary inversion: $Q-b_{\chi}-bN \rightarrow V-aN-b_{\chi}$):
Are there Really Gaps in Filler-gap Constructions?

Note that categorial grammar account does not predict gaps per se. Rather arguments are predicted to be satisfied at beginning of verb phrase.

This is not a problem for Gibson, but it raises an interesting question…

Pickering & Traxler ’99: Do gaps exist or is categorial explanation better?

- **stimuli:** sentences in self-paced reading
  - (a) ‘That’s the pistol with which \( i \) the killer shot the helpless man \( \epsilon_i \).’
  - (b) *‘That’s the pistol in which \( i \) the killer shot the helpless man \( \epsilon_i \).’

- **measure:** self-paced reading durations

- **results:** large delay at ‘shot’ in ungrammatical sentence (b).

Results support categorial explanation: dependencies go directly to verb.
For next time... 

Have a good break!

Read:

- Traxler ch 5, pp. 187–210