Let's assume for the purposes of today's discussion that John Ross' (1967) observation that rightward movement is 'upward bounded' by a clausal boundary is a roughly correct empirical generalization. From our point of view, this raises the question of what is independently known about how comprehenders treat the clause in syntactic comprehension (or production). Today we're simply going to ask 'how do clausal boundaries impact the availability of syntactic information during parsing?' We are going to focus simply on establishing the empirical facts of the matter, and we'll explore the explanations for these facts in the weeks that follow.

1 Evidence from recognition and recall

Common paradigms for evaluating the robustness of memory encodings come in two flavors: recall tasks and recognition tasks. A recall task involves presenting participants with a linguistic stimulus during a 'study' phase, where they form a memory encoding of the stimulus. This may be done in an 'explicit' or 'implicit' manner. After study, participants are subsequently asked to recall the stimuli at some delay (during the 'test' phase). Success in the test phase depends on forming a robust and reliable encoding of the stimulus during study, and maintaining that representation over some 'retention interval.'

A recognition experiment has the same study-retention-test structure, but adopts a different strategy for the test phase. In recognition tasks, the test consists of presenting a linguistic stimulus, and asking the participant to evaluate whether the stimulus was present in the study phase. Here the participant is asked to discriminate using a forced choice task targets (items presented at study) from lures (items that were not presented at study).

1.1 Sachs (1967)

Sachs (1967) used a recognition task to study participants’ memory for linguistic stimuli. Participants read short passages, each of which contained a target sentence. The test phase was cued by a bell played on tape, and occurred immediately after the target sentence, 80 syllables after the target sentence, or 160 syllables after the target sentence. At test, Sachs tested recognition memory for the target sentence (correct recognition, or hits), as well as the amount of confusion (incorrect recognition, or false alarms) caused by a number of different types of lure:

1. Target: He sent a letter about it to Galileo, the great Italian scientist.
2. Syntactic change: He sent Galileo, the great Italian scientist, a letter about it.
3. Semantic change: Galileo, the great Italian scientist, sent him a letter about it.
4. Passivization: A letter about it was sent to Galileo, the great Italian scientist.

Figure 1: Data from Sachs (1967); accuracy on recognition task.
This pattern of results suggests that participants jettison the surface syntactic encoding of a sentence somewhere between 0 and 80 syllables after comprehending that sentence. This is sometimes referred to as 'gist' memory: we retain robust encodings of some aspects of the sentence's meaning, but apparently very little about its actual form. Similar results are observed in recall paradigms see Potter and Lombardi (1990), Potter (1993), and Potter and Lombardi (1998).

1.2 Jarvella (1971)

Jarvella (1971) used an 'interrupt-and-recall' paradigm. Participants listened to audio recordings of sentences, and were periodically interrupted by the experimenter and asked to write down, verbatim, the stretch of the passage since the last test pause. Jarvella tested recall for two types of configurations:

1. Long: Context (7 words), Previous (6 words), Immediate (7 words).
2. Short: Context (7 words), Previous (6 words), Immediate (7 words).

The results:

![Table 1: Mean score for each test type](image)

<table>
<thead>
<tr>
<th>Syllables of Interpolated Material</th>
<th>0</th>
<th>60</th>
<th>180</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>3.72</td>
<td>2.71</td>
<td>2.47</td>
</tr>
<tr>
<td>P/A</td>
<td>3.87</td>
<td>.96</td>
<td>.78</td>
</tr>
<tr>
<td>Form</td>
<td>2.22</td>
<td>.55</td>
<td>.17</td>
</tr>
<tr>
<td>Ident</td>
<td>3.44</td>
<td>.44</td>
<td>.95</td>
</tr>
</tbody>
</table>

**Figure 2:** Data from Sachs (1967); accuracy scores weighted by 1-5 confidence rating. Positive is accurate, negative is inaccurate, 0 is chance.

This graph plots the probability of correct recall as a function of serial position in the word string from the end of the sentence. Jarvella calculated $P(\text{correct})$ in two different ways. Closed circles are standard free recall accuracy: a serial position is counted as correctly recalled if the word that was in that position in the target sentence (or list) is recalled, regardless of its context. The dark circles are $P(\text{correct})$ based on a running memory span: a word is correctly recalled only if all subsequent words are also correctly recalled.

On the assumption that a robust syntactic encoding helps to support correct recall (particularly of the running memory span), these data suggest the following conclusions. One, crossing a sentence boundary leads to a dramatic drop in the robustness of a syntactic encoding. Two, crossing a boundary between a non-finite adjunct clause and its matrix sentence leads to a smaller, but reliable drop in how well represented a syntactic encoding is.

Experiment 2 addressed the concern that the curves shown above reflect a response strategy on the part of the participants. In Experiment 1, they were not explicitly told to recall the sentences in left-to-right order, leading to the possibility that they started with the local clause before attempting to recall more distant clauses, which could result in an apparent clause-boundary effect in accuracy. Experiment 2 adopted a cued recall paradigm, and gave participants a cue word.
to prompt recall. Their task was to write down the words they recalled in a strict left to right order from the position of the cue word to the end of the immediately preceding clause. This technique replicated the effects we saw above.

Some further interesting observations about the free recall tasks and linguistic memory. Tyler and Marslen-Wilson (1978) performed a similar task with children of varying age groups, and showed that 7 and 11 year olds pattern like Jarvella’s undergrads in that they demonstrate perfect recall for the immediate clause in an interrupt-and-recall paradigm, with a sharp drop beyond the immediate sentence. However, younger children (5 year olds) did not show such a sharp clausal boundary effect, most dramatically when the immediate clause was a subordinate clause.

A second observation about Jarvella’s data is the ‘bowed’ shape of the recall curves for clauses beyond the immediate clause. This is of interest because it has the same functional form as the free recall curves observed for lists of unrelated words. The classic demonstration of this bowed serial position curve was Bennet B. Murdock (1962), who asked participants to remember lists of words of varying lengths and prompted them to recall elements of that list at test. His results:

![Figure 4: Data from Bennet B. Murdock (1962)](image)

Murdock found that presentation speed and list of length did not change the qualitative pattern of results: recall in all cases seems to have a bowed shape that demonstrates a primacy effect and a recency effect. Items at edges have more robust memory encodings, and subsequent research suggests that these two effects have different underlying sources.

1.3 Caplan (1972)

Caplan (1972) used a probe recognition task to build a case for clause boundaries in perception. Like Jarvella’s interrupt-and-recall paradigm, Caplan had participants listen to short audio passages. At the end of each sentence, participants heard a probe word, and were asked to indicate whether the probe was in the immediately preceding sentence or not. Reaction times were recorded Caplan controlled for linear / temporal position of the target word vis-a-vis the probe:

![Figure 5: Materials from Caplan (1972) Experiments 1 and 2](image)
Experiment 1 used audio presentation of the probe; Experiment 2 used a visual presentation of the probe word. One important reason to test both audio and visual presentation of the probe word is to ensure that participants are using a modality independent representation (by hypothesis, a parse) to complete the task. Otherwise, the task simply measures the decay of information in a modality-specific representation. The probe word was presented 100ms after the offset of the sentence:

Figure 6: Materials from Caplan (1972) Experiments 1 and 2

Experiments 1 and 2 showed that the clausal organization, rather than linear distance measured in syllables, is what controlled probe recognition latencies. On average, responses to probes in the recent sentence were 30ms faster for audio probes, and 60ms faster for visual probes. The number of syllables did not in itself impact recognition latencies (no main effect of syllable number).

However, he did not control for intonational factors, nor did he control for lexical identity in Experiments 1 or 2, raising the possibility that the results here are spurious. That is, they may reflect an accidental correlation between, say, lexical frequency and clausal position. Experiments 3 and 4 adopted a within-items design to rectify this concern:

(4) a. Before final clause: Now that artists are working in oil, prints are rare.
   b. In final clause: Now that artists are working fewer hours, oil prints are rare.

He divided sentences in (4) into two lists (i.e. Latin Square) and presented them to two groups. The last five words (i.e. 'oil prints are rare') were spliced from the 'in final clause' conditions into the 'before final clause' conditions, controlling for acoustic cues. Naive raters could not distinguish spliced from natural examples.
In all cases the results are in agreement: recognition latencies are lower for probes that are contained in the immediately preceding clause. This effect is not reducible to modality-specific representations, nor syllabic or temporal distance, nor acoustic cues. Instead, there appears to be a difference in the availability (or accessibility) of lexical information across a sentence boundary.

1.4 Chang (1980)

Chang (1980) presents an interesting follow-up to Caplan (1972). Rather than using auditory presentation, Chang presented target sentences to participants using visual self-paced reading. Participants read sentences in a word-by-word fashion, with a visual probe immediately (i.e. oms ISI) following the sentence (distinguished by a 2 degree offset on the screen). Additionally, Chang asked participants a comprehension question following the critical probe word, to ensure that participants did not rely on non-linguistic representations of the sentences to perform the task. He tested ‘natural’ and ‘controlled’ sentences; his ‘controlled’ sentences were drawn from Caplan’s materials for Experiments 3 and 4. The ‘natural’ sentences were drawn from Reader’s Digest and other esteemed publications:

5

(5)  a. Before final clause: When the wind began to get stronger in the early afternoon the sun warmed us.
  b. In final clause: As midsummer nears and the pond begins to shrink frogs crawl into the mud.

Chang adopted the natural materials to ensure the generalizability of the results, and also claims that they suffered from fewer ambiguities than Caplan’s materials (although given this concern, his parade examples in (6) are somewhat poorly chosen!). His results mirror Caplan’s, and extend them by demonstrating that clausal position affects recognition accuracy:

Experimental procedures were the same as in Experiment 1. Chang adopted this design to ask whether establishing pronominal reference across clauses was enough to reactivate information in a distant clause. In the reinstated pronoun condition, the lexical item John is in the distant clause, but it is linked to the most recent clause via a referential link. Chang asked if this link would serve to increase the availability (or accessibility) of a distant linguistic encoding.
The reinstated-control contrast replicates Chang’s Experiment 1, and confirms the results in Caplan (1972). The RT pattern in Experiment 2 confirms a main effect of reinstatement, qualified by an interaction between reinstatement and pronominalization. Resolving this interaction with planned pairwise comparisons revealed that the reinstated pronoun condition was faster (and more accurate) than either of the control conditions. The establishment of a referential link appears to ‘refresh’ the encoding of a distinct lexical item, making it more available/accessible. The empirical finding of ‘refreshing’ effect will provide a helpful constraint on the development of theoretical models of the clause bounded effect supported by the recognition results seen so far.

1.5 Gernsbacher et al. (1989)

Another interesting set of probe recognition studies comes from Gernsbacher et al. (1989). They note that in their work, they observe a robust advantage for elements that are ‘first mentioned.’ The first-mention advantage is analogous to Murdock’s primacy effect, and refers to an increase in availability/accessibility for the left-most participant (e.g. first DP) in a sentence. For instance, compare the following two sentences:

(7) a. Tina beat Lisa in the state tennis match.
   b. Lisa beat Tina in the state tennis match.

Recognition latencies for Tina had been observed to be faster when it occupies the matrix subject position. Gernsbacher et al. (1989) hypothesized that that the first-mentioned participant serves as the starting point for syntactic parsing and interpretative processing, and so receives special prominence. They present this as an apparent contradiction of Caplan (1972), which they interpret as a sort of recency effect. They hypothesized that multi-clausal structures might be treated as distinct parsing domains, such that when one clause is embedded inside another, what counts as ‘first-mentioned’ is evaluated within the local clause. She predicted, based on this, that one should be able to observe both a locality bias and a first mention bias depending on which clause the comprehender is attending to upon presentation of a probe.

They used visual RSVP presentation (300ms/word, plus 16ms per character, 150ms ISI) to present sentences to participants, presented a visual probe, and followed this with a comprehension question. Across experiments, they manipulated the latency of the presentation of the probe word. Probe words were either immediately following the presentation of the sentence (Experiment 1), 150ms following presentation of the sentence (Experiment 2), 1400ms following presentation of the sentences (Experiment 3) or 2000ms after the presentation of the sentence (Experiment 4). Their materials consisted of transitive sentences with a finite subordinate temporal clause, which was either in initial or final position, or of conjoined sentences (with both clause orders):

(8) a. As Lisa set up the tent, Tina gathered the kindling.
   b. Lisa set up the tent as Tina gathered the kindling.
   c. Lisa set up the tent, and Tina gathered the kindling.
Probe words were either to the first name in the conjunct or the second, and either from the local or distant clause. Experiment 5 tested this with a brief probe latency, and Experiment 6 tested this with 2000ms probe latency.

2 Evidence from processing agreement and binding dependencies

In addition to evidence from tasks that explicitly target memory, evidence from locality effects in agreement interference and processing referential expressions points to a clause-boundedness effect in syntactic comprehension.

2.1 Bock and Cutting (1992)

One piece of implicit evidence for clause-boundedness comes from agreement attraction effects in comprehension. Bock and Miller (1991) designed a cued production task to elicit agreement attraction errors in the lab. They had participants listen to a sentence preamble (a subject phrase) over headphones, and then asked the participants to complete the sentence any way they saw fit. They manipulated the morphological features on each of the nouns in a complex subject. The crucial comparison involved the number features associated with an embedded, distractor noun:

(10)  a. No distractor preamble: The key to the cabinet ...
    b. Local distractor preamble: The key to the cabinets ...

Bock and Miller famously found that the presence of a local distractor caused an increased amount of erroneous plural verb production in (10-b) compared to (10-a), suggesting that non-subject, syntactically inaccessible nouns can interfere with the construction of an agreement dependency. This effect came to be known as the agreement attraction effect, which is used descriptively to refer to instances in which a non-subject distractor DP appears to control agreement on a verb or pronoun.

Of interest for the current purposes is the relative degree of interference engendered by distractors in different syntactic positions. Bock and Cutting (1992) explored this explicitly. They investigated three types of preambles (collapsing across their Experiments 1 and 2):

(11)  a. Phrase preamble: The editor of the history book(s) ...
    b. Relative clause preamble: The editor who rejected the book(s) ...
    c. Complement clause preamble: The report that they controlled the forest fire(s) ...

Their results:
Bock and Miller found that the rate of erroneous plural production dropped when the distractor was in a different clause than the target head noun (Experiment 1: .47 for phrase preambles to .20 for RC preambles; Experiment 2: .15 for phrase preambles to .07 for complement clause preambles; Experiment 3: .15 for phrase preambles to .06 for complement clause preambles).

Franck et al. (2010) used a similar cued production task to measure the rate of agreement errors in French. Participants studied a sentence preamble, and were presented with an infinite verb. They were instructed to complete the sentence with the capitalized infinite form. All verbs used had a phonological distinction between singular and plural third person (e.g. sort versus sortent). They presented a range of experiments, but one in particular provides another data point for our discussion. They contrasted the production of verbs in relative clauses with those in complement clauses, manipulating the number of the DP that precedes the critical clause:
Although the complement and relative clause examples were nearly string identical, the critical distractor noun *la prisonnière* is syntactically active only in the relative clause. The results:

<table>
<thead>
<tr>
<th>Agreement errors</th>
<th>Complement clause</th>
<th>Relative clause</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>1.4 (.37)</td>
<td>2.8 (.16)</td>
<td>2.1 (.13)</td>
</tr>
<tr>
<td>Mismatch</td>
<td>2.9 (.16)</td>
<td>11.1 (.31)</td>
<td>6.9 (.25)</td>
</tr>
<tr>
<td>Total</td>
<td>5.9 (.13)</td>
<td>14.9 (.25)</td>
<td>4.3 (.20)</td>
</tr>
</tbody>
</table>

Foraker used self-paced reading with a phrase-by-phrase moving window (indicated by brackets in first example). Reading times on each region were measured. The results:

2.3 *Dillon* (2011); *Foraker* (2003)

Another source of evidence about the primacy of the local clause in computing online syntactic dependencies comes from studies on the processing of bound anaphora. We'll look at two studies that bear on this. The first is *Foraker* (2003). She studied the processing of reflexives in binding-theory exempt positions (informally referred to as *logophors*). She looked at sentences containing reflexive pronouns in coordination environments such as the following:

\[ (12) \text{ a. First-mentioned antecedent: } Megan \text{ wondered if Isaac had found out that } Megan \text{ wanted to invite Sally and she herself to the birthday party.} \]

\[ (13) \text{ a. First-mentioned antecedent: } Albert \text{ was upset when Debbie didn't care that Rachel had endangered Gordon and him on the climbing trip.} \]

\[ (13) \text{ b. Middle antecedent: Debbie was upset when } Albert \text{ didn't care that Rachel had endangered Gordon and him on the climbing trip.} \]

\[ (13) \text{ c. Most recent antecedent: } (\#?-BWD) \text{ Rachel was upset when Debbie didn't care that } Albert \text{ had endangered Gordon and him on the climbing trip.} \]

She additionally manipulated the identity of the pronominal element, and looked similar antecedent configurations when the coordinated anaphor was a pronoun:

\[ (13) \text{ a. First-mentioned antecedent: } Megan \text{ wondered if Isaac had found out that } Megan \text{ wanted to invite Sally and herself to the birthday party.} \]

\[ (13) \text{ b. Middle antecedent: Isaac wondered if } Megan \text{ had found out that Rick wanted to invite Sally and herself to the birthday party.} \]

\[ (13) \text{ c. Most recent antecedent: Rick wondered if Isaac had found out that } Megan \text{ wanted to invite Sally and herself to the birthday party.} \]
reverse was true of pronouns (perhaps due to grammatical constraints penalizing coreference in that configuration).

Dillon (2011) examined the processing of the long-distance reflexive *ziji* in Mandarin Chinese. *ziji* can be bound in its local clause, or by subjects in dominating clauses (Huang, Li, & Li, 2009). We examined the following three types of sentences:

(14) a. Long-distance antecedent: *Chushi* baogao you duo cengjing tangshang-le *ziji*, suoyi cizhi-le.
    Chef say deep-fryer scalded *ziji*, so (pro) resigned.

b. Local antecedent: Yiliao baogao *chushi* cengjing tangshang-le *ziji*, suoyi cizhi-le.
   Medical report say chef scalded *ziji*, so (pro) resigned.

c. No antecedent: *Yiliao baogao you duo cengjing tangshang-le *ziji*, suoyi cizhi-le.
   Medical report say chef scalded *ziji*, so (pro) resigned.

We recorded event-related potentials and collected acceptability judgments for these three sentences. Acceptability results showed that the presence of a c-commanding, animate antecedent led to fairly high rates of endorsement for long-distance (81%) and local (73%) binding conditions; the normatively ungrammatical no antecedent condition was endorsed at a fairly high rate of (51%). Post-hoc debriefings and questionnaire studies suggest that the preferred interpretation in this case is that *ziji* refers to the implicit author of e.g. the medical report. Note that the offline acceptability data suggests that participants prefer *ziji* to be long-distance bound.

Although the offline results suggest a preference for long-distance construals of *ziji*, the online ERP data suggest a different pattern.

3 Where we’re at

Results from a range of methodologies and a range of linguistic structures indicate that syntactic information in the local clause is more available than information outside the local clause. Specifically we saw evidence from:

- Recall: recall of syntactic form is only preserved for the local clause.
- Recognition: recognition of probe words is fastest, most accurate for the local clause.
• Anaphor resolution: for anaphors that may be bound locally or long distance, there is greater processing effort required to construct long-distance binding.

• Agreement attraction: interference effects on agreement processing are strong for distractors within the local clause.

References


