Chapter 6
Basic Evidence for the Model

This chapter presents facts that bear specifically on the comprehension model that embeds a grammar in an analysis-by-synthesis architecture. Various other models can account for some of the individual facts that we consider here, but the facts in aggregate uniquely support the architecture we are proposing.

The chapter is organized in terms of the approximate progression of activities that occur during comprehension. We first consider evidence on the initial application of low-level syntactic structure, which we call "pseudosyntax." Then we consider evidence on early behavioral access to meaning, followed by evidence on late behavioral access to "real syntax." We conclude with a discussion of the model's explanation of rapid judgments about sentence acceptability.

6.1 Pseudosyntax

Pseudosyntax consists of the immediate initial processes that isolate major phrases, differentiate lexical categories, and assign initial thematic relations. Pseudosyntax involves recognition of function morphemes and lexical categories, which segregate and distinguish phrases and verbs. Assignment of words to syntactic categories and major phrases coincides with the application of frequent sentence patterns that assign these phrases to thematic roles. The sentence patterns that are appropriate for a particular sentence depend on subcategorization properties of verbs. The review of other comprehension models in chapter 4 highlighted several kinds of syntactically relevant information that can be extracted from statistically supported patterns. For example, Juliano and Bever (1988) verified that a simple connectionist model with a moving three-word window can learn to predict utterance boundaries and, thereby, phrase boundaries. Mintz (1997) showed that a three-word window provides sufficient information to differentiate the lexical category of the middle word. Many models invoke the priority and cue validity of the statistically dominant thematic pattern: NVN = actor action patient. Each of these investigations with formal models

Box 6.1
The monitoring task

In the word-monitoring paradigm, subjects receive a cue word to listen for. A stimulus sentence or sentences is then presented. In critical cases, the stimulus material contains an instance of the cue word. This instance of the cue word in the stimulus material is called the *target*. For example:

Cue word: ONLY
Stimulus: "The owner rented only to older couples"
Target word: only

The measure of performance is either accuracy or speed of detecting the target word. If speed is measured, a timing tone is usually placed on tape to coincide with the target word. The timing tone triggers a timer, which stops when the subject makes a response. The subject's response may be a key press or a vocal response. It is assumed that as response times decrease, the amount of effort needed to recognize the target word decreases.

Monitoring tasks have used a variety of other targets. In many studies, subjects monitor for a particular phoneme, such as /k/. Sometimes the target is a synonym of a cue word, such as a synonym of *old*. Subjects sometimes monitor for category label targets, such as an "age descriptor," or for nonsense word targets ("respond when you hear a nonsense word"). Box 9.1 describes a monitoring task in which the subject listens for a change of speaker (e.g., "respond when you hear a word spoken by someone of another gender"). See Foss 1969; Marslen-Wilson and Tyler 1980; Marslen-Wilson, Tyler, and Seidenberg 1978.

supports the claim that surface sequence cues and patterns provide a great deal of reliable information relevant for initial syntactic analysis.

6.1.1 Early Access to Function Words

Friederici (1985) demonstrated that comprehenders access function words more rapidly than content words. Subjects monitored two-sentence texts for a particular target word. The two sentences were semantically unrelated. The target word was either a function word (article, demonstrative, or quantifier) or a content word (noun or adjective). After hearing the target word, subjects heard the text material, and pressed a response key as soon as they recognized the target word in text (see box 6.1). Some example materials appear below (with target word italicized):

Function-word target, unrelated context

(1) Die Kommandos beim Wenden eines Segelbootes sind oft sehr knapp.
the commandos in-the putting-about of a sailboat are often very close

(2) Der Besitzer vermietet *nur* an ältere Ehepaare.
the owner rented only to elderly couples

Table 6.1
Mean word-monitoring times (ms) for function words and content words in short discourses

Target	Related	Unrelated
Function word	319	328
Content word	338	369

Source: Adapted from Friederici 1985, table 1

Function-word target, related context

(3) Die Auflagen beim Anmieten einer Wohnung sind oft sehr strikt.
the restrictions on-the renting of-a home are often very strict

(4) Der Besitzer vermietet *nur* an ältere Ehepaare.
the owner rented only to elderly couples

Content-word target, unrelated context

(5) Der verliebte Student entschloss sich, ins Grüne zu fahren.
the lovesick student decided himself into-the country to drive

(6) Der Mann hoffte, *Geld* zu gewinnen.
the man hoped money to win

Content-word target, related context

(7) Der verarmte Spieler entschloss sich, ins Kasino zu gehen.
the impoverished player decided himself in-the casino to go

(8) Der Mann hoffte, *Geld* zu gewinnen.
the man hoped money to win

Response times were faster for function words than for content words, especially for unrelated contexts (table 6.1). A context sentence semantically related to the second sentence reduced response times by 31 ms for content-word targets but only 9 ms for function-word targets. In addition, semantic context reduced monitoring times for function words that carried optional information unrelated to the main-verb arguments, such as prepositions that introduce locations. Semantic context did not affect monitoring time for function words that carried structural information, such as prepositions that introduce obligatory arguments.

6.1.2 Function Words and Low-Level Phrases

Accessing function words early could enable comprehenders to determine quickly the words that fill major thematic roles. If one hears *the*, one knows to search for a head noun. In fact, O'Seaghdha (1997) has shown that determiners prime nouns. He used a cross-modal naming task in which subjects heard sentence fragments followed by a

Box 6.2
The cross-modal naming task

The cross-modal naming task requires subjects to "name" or say aloud a visually presented word while listening to speech. The task is called *cross-modal* because two modalities are used: the stimulus is presented in speech, but the word to be named (i.e., the target) is presented in print. The spoken stimulus may end when the target appears, or it may continue. For example,

Stimulus (speech): "The message of that"
Target (print): LETTER

Typically, the relatedness of the target to the linguistic stimulus is varied. Speed of naming the target is measured by starting a timer at the point at which the target word appears. It is assumed that faster response times indicate that less processing effort is needed to recognize the target word because of expectations induced by the linguistic material. See Tyler and Marslen-Wilson 1977.

target word to say aloud (see box 6.2). The sentence fragments were consistent with the next word being either a noun or a verb. Some sample materials appear here:

(9) *Grammatically consistent, semantically related*
   a. the message of that LETTER
   b. the message that was SENT

(10) *Grammatically consistent, semantically unrelated*
   a. the nose and the LETTER
   b. the nose that he SENT

(11) *Grammatically inconsistent, semantically related*
   a. the message of that SENT
   b. the message that was LETTER

(12) *Grammatically inconsistent, semantically unrelated*
   a. the nose and the SENT
   b. the nose that he LETTER

The target word was either consistent or inconsistent with the syntactic context in the sentence fragment. For example, *the message for that* is consistent with a noun target, while *the message that was* is consistent with a verb target. The target word was also either related to a content word in the sentence fragment (*message-letter*, *message-sent*) or not related. The sentence fragments consistent with a following noun ended with determiners (*the*, *a*, *this*, *that*, *her*, *many*, *its*, *those*) and in some cases they ended with prepositions (*of*, *from*, *in*). The sentence fragments consistent with a following verb ended with pronouns or proper names (*she*, *they*, *he*, *John*), modal verbs (*would*,

Table 6.2
Mean naming times (ms) for noun versus verb targets

	Context bias		
Target	Noun	Verb	Bias effect
Related			
Noun	470	492	22
Verb	498	494	4
Unrelated			
Noun	480	499	19
Verb	505	498	7

Source: Adapted from O'Seaghdha 1997, fig. 2

should, *could*), or adverbs (*just*, *often*, *first*). In fact, the sentence fragments above are only "disposed" to a following noun or verb. For example, a noun-disposing context *the message of that* can precede a noun or an adjective (e.g., *the message of that long letter*). Similarly, a verb-disposing context like *the message that was* can precede a passive participle (*sent*), an auxiliary verb (*being*), an adverb (*quickly*), an adjective (*long*), a determiner (*a mistake*), or even a noun (*ear-piercing*). On this basis, the noun contexts were more constraining than the verb contexts.

Response times to name the target were faster when the syntactic category of the target was consistent with the syntactic context (table 6.2). This effect was stronger for noun than verb targets. The syntactic context effect was similar for semantically related and semantically unrelated targets. These results demonstrate that comprehenders attend to function words, and these can help them identify nouns in particular.

6.1.3 Number Morphemes

Pseudosyntactic structures are assigned by templates triggered by superficial cues. Attending to explicitly recognizable function morphemes attached to words can also assist in comprehension (Friederici 1985). Plural number on the end of a noun is such a function morpheme. In English, singular nouns are unmarked, but plural nouns are marked, as in *speech-speeches*, *author-authors*, and so on. Therefore, we can expect that a plural morpheme on a noun will lead the comprehender to expect plural marking on a following verb. At the same time the unmarked singular ending does not trigger a local pattern as strongly.

Behavioral evidence for comprehenders' sensitivity to plural morphemes on nouns comes from Nicol, Forster, and Veres (1997). In a subject- or self-paced whole-sentence reading task (see box 6.3), they varied factorially the number marking on two preverbal nouns and the verb. A portion of their design and materials appears below:

Box 6.3
The subject-paced reading task

In a subject-paced or self-paced reading task, linguistic material is presented segment by segment on a computer screen. The subject reads the material at his or her own pace by pressing a key on the keyboard. Each key press removes the previous segment, measures and records the amount of time the subject spent on the previous segment, and displays the next segment. The display unit may be any size. The most frequently used display units are words, phrases, and whole sentences (see box 6.4 for more details about the word-reading task). These three display units are illustrated below with a separate line for successive displays:

Word:
The
army
relentlessly
attacked
the
city.

Phrase:
The army
relentlessly attacked
the city.

Sentence: The army relentlessly attacked the city.

Sometimes a single sentence is displayed, and the subject's task then is to press a key when the subject feels that he or she understands the sentence. Response times decrease as processing load decreases. See Aaronson and Scarborough 1976; Just, Carpenter, and Woolley 1982.

(13) *Singular head noun, singular local noun*
The author of the speech was subsequently well rewarded.

(14) *Singular head noun, plural local noun*
The author of the speeches was subsequently well rewarded.

(15) *Plural head noun, singular local noun*
The authors of the speech were subsequently well rewarded.

(16) *Plural head noun, plural local noun*
The authors of the speeches were subsequently well rewarded.

Table 6.3 shows that when the head noun was singular (*author*), reading times were longer when the local noun was marked as plural (2044 ms) than when it was singular (1920 ms). When the head noun was plural (*authors*), there was only a 1-ms difference in reading time depending on the local noun. The failure to find elevated reading times when the local noun is singular and the head noun is plural demonstrates the early use of overt number information, which appears on plural nouns but not on

Table 6.3
Mean sentence-reading times (ms) depending on number of head noun and local noun

	Head noun	
Local noun	Singular	Plural
Singular	1920	1994
Plural	2044	1995
Difference (P−S)	124	1

Source: Adapted from Nicol, Forster, and Veres 1997, table 2

singular nouns. Nicol and colleagues replicated these results with a two-choice word-by-word selection task (see section 6.4.2).

In the study by Nicol, Forster, and Veres, the explicit marker to plural number on a local noun primes plural marking on the following verb. Reading times may be elevated when the head noun is singular and the local noun is plural because the comprehension system incorrectly assigns the local noun as the subject of the verb. This simple result reflects the power of local syntactic patterns in assigning an initial pseudosyntax.

6.1.4 NVN Is a Canonical Word Order in Pseudosyntax

Further evidence for early pseudosyntax comes from a large body of research on the comprehension of sentences that violate the canonical pattern NVN = agent-action-patient (see chapter 2). We discuss the recent literature in detail in chapters 7 to 9 when we consider various applications of the model to comprehension and acquisition. In this section, we present some highlights of this research.

The purpose of isolating words into major phrases is to fill thematic positions. The most common pattern of thematic roles in English sentences is the NVN pattern. In this pattern, the initial noun phrase is the agent, the verb corresponds to the action that the agent performed, and the first postverbal noun phrase is the patient or theme of the action. The difficulty of sentences with a surface order of phrases not in the canonical pattern is evidence for the early application of the NVN template.

A well-attested fact in psycholinguistics is that a passive sentence is harder to process than the corresponding active. Thus, in matching sentences to pictures, it is harder to match a passive sentence than an active (Slobin 1966):

(17) *Active*
The dog is chasing the cat.

(18) *Passive*
The cat is being chased by the dog.

In judging the grammaticality of sentences, passives are harder than actives (Forster and Olbrei 1974). In measures of sentence reading time, passive sentences take longer than actives as in (19) and (20), which are adapted from Davison and Lutz (1985):

(19) *Active*
A man in California raised a twenty-five-year-old dachshund.

(20) *Passive*
A 25-year-old dachshund was raised by a man in California.

Similar processing differences appear in other pairs of canonical/noncanonical sentences. For example, children have more difficulty acting out object-cleft sentences than subject-cleft sentences (Bever 1970a). Object-cleft sentences violate the canonical order.

(21) *Subject cleft*
It was the dog that chased the cat.

(22) *Object cleft*
It was the cat that the dog chased.

Memory load, as measured by difficulty of recall of unrelated names presented during the sentence, is greater when processing object relative clauses (Wanner and Maratsos 1978), which violate the canonical order:

(23) *Subject relative with interpolated word list*
The witch who despised JOHN GEORGE SAM BILL HANK sorcerers frightened little children.

(24) *Object relative with interpolated word list*
The witch whom sorcerers JOHN GEORGE SAM BILL HANK despised frightened little children.

The syntax that is assigned initially in each of these cases is rudimentary. This pseudosyntax is based on simple matching of major phrases to the canonical sentence patterns in order to assign argument roles. For example, in a subject relative clause the phrases match up with the canonical template as follows:

who	=	initial NP	=	agent
despised	=	verb	=	action
sorcerers	=	second NP	=	patient

An object relative clause is more difficult because its phrases do not match the canonical sentence pattern:

whom	=	initial NP	=	patient
sorcerers	=	second NP	=	agent
despised	=	verb	=	action

Argument roles are assigned as a concomitant of the matching of major phrases to common sentence patterns.

6.1.5 Syntactic Categories

The *syntactic category* of a word refers to its categorical role in a sentence. These roles include noun, verb, adjective, and so on. In the Late Assignment of Syntax Theory (LAST), syntactic category is assigned as a concomitant of the organization of words into major phrases that fill thematic roles. Homographs such as *duck*, *train*, and *fire* are ambiguous with regard to syntactic category. LAST predicts that if a word can fill a thematic role, it is assigned to the corresponding syntactic category. Assignment of a word to a needed syntactic category, however, will be harder if the word is rarely used as that category.

Boland (1997b) examined the influence of a variety of probabilistic factors in the initial assignment of syntactic category. She studied sentences that contained a homograph like *duck*:

(25) *Possessive pronoun, noun*
She saw her duck and chickens near the barn.

(26) *Accusative pronoun, verb*
She saw her duck and stumble near the barn.

In these sentences, the appropriate syntactic category of *duck* as a noun or a verb cannot be fully determined until the sixth word (*chickens* or *stumble*). The frequency of use of a word as a noun or verb influences its immediate assignment. For example, *duck* is used more often as a noun than as a verb and tends to be assigned initially the category of noun, all other factors equal. In actual sentences other factors, such as local context, are usually not equal. Based on O'Seadgha's (1997) results, we would expect that when the preceding pronoun is possessive, *duck* will be assigned the category of noun, and when the pronoun is accusative, *duck* will be assigned the category of verb:

(27) *Preceding possessive pronoun*
She saw his duck.

(28) *Preceding accusative pronoun*
She saw him duck.

Since *her* may be either possessive or accusative, it does not establish with certainty the category of *duck*. However, we can expect that the frequency with which *her* is used as a possessive versus accusative will influence its initial case assignment as well, and also influence the immediate assignment of *duck*.

Another potential factor in assigning a syntactic category is the ambiguity of the main verb *saw*. This verb may take a noun phrase object as a transitive verb, or a

sentential complement as an object control verb, as in:

(29) *Noun phrase object*
She saw her.

(30) *Object with sentential complement*
She saw her stumble.

If the main verb subcategorizes only for a noun phrase object, *her* will initially be assigned the accusative case, yielding an initial meaning that corresponds to (31).

(31) She hit her.

(32) *She hit her stumble.

Other verbs, such as *think*, however, require additional cases depending on the subcategorization frame, as in:

(33) *Sentential complement*
I think she left.

(34) *Noun phrase object*
I think her thoughts.

(35) *Object with sentential complement*
I think her a fool.

(36) *Pronoun object*
*I think her.

The possible subcategorization frames and their frequency of use can influence the assignment of syntactic category to a homograph as well. Boland used a moving-window reading task to examine the effects of frequency of use of a homograph and discourse context on the initial assignment of words to syntactic categories (box 6.4).

The target homograph appeared in sentences such as one of the following:

(37) *Unambiguous possessive pronoun, noun*
She saw his duck and chickens near the barn.

(38) *Unambiguous accusative pronoun, verb*
She saw him duck and stumble near the barn.

The target sentences were preceded by a context sentence such as the following:

(39) *Context bias toward possessive pronoun, noun*
As they walked around, Kate looked at all of Jimmy's pets.

(40) *Context bias toward accusative pronoun, verb*
As they walked around, Kate watched everything that Jimmy did.

Box 6.4
The moving-window reading task

The moving-window reading task is a variant of subject-paced reading, discussed in box 6.3. As discussed in box 6.3, single-word presentation often displays one word at a time in the center of the screen. In the moving-window paradigm, each screen contains information about punctuation, spacing, and the length of the words, but the letters of only one word are displayed. For example, the lines below represent successive displays:

The **** ********** ******** *** ****.
*** army ********** ******** *** ****.
*** **** relentlessly ******** *** ****.
*** **** ********** attacked *** ****.
*** **** ********** ******** the ****.
*** **** ********** ******** *** city.

Each key press replaces the previously displayed word with asterisks, records the time that the previous unit was displayed, and shows the next word together with asterisks corresponding to the remaining words on the line. The amount of time the subject spends reading a word is taken to indicate the processing load needed for that particular word. See Just, Carpenter, and Woolley 1982.

The first discourse context highlights Jimmy's possessions, while the second highlights things Jimmy did. If discourse context influences the initial assignment of syntactic category, reading times for the pronoun *his* ought to be faster when the discourse context is biased toward the possessive, and reading times for *him* ought to be faster when it is biased toward the accusative. If syntactic category assignment depends on frequency of use of the noun-versus-verb meanings of the homograph, reading times in possessive contexts ought to decrease as the noun meaning of the homograph increases in frequency.

Discourse context did not have a significant effect on reading times on any word (see table 6.4), though its effect approached significance on the disambiguating word (*stumble*, *chickens*). This result indicates that discourse context does not play a role in the immediate assignment of syntactic category.

On the other hand, Boland did find that homograph bias was related to reading time. Contingent homograph bias was determined by asking subjects to complete fragments such as *she saw her duck* ... Homographs that more often produced completions with a noun usage of the homograph are "noun biased," while those that produced more completions with verb usage are "verb biased." The size of the difference between noun and verb completions indicates the strength of homograph bias. Boland found a significant correlation ($r = .75$) between strength of homograph bias and reading times for the word following the homograph (i.e., *and*). When the pronoun was the accusative *him*, which allows a following verb, reading times for *and*

Table 6.4
Mean word-reading times (ms) depending on discourse context and local context

	Word			
	him/his	duck	and	stumble/chickens
Him				
Poss discourse	340	350	385	425
Acc discourse	345	355	350	405
Discourse effect	−5	−5	35	20
His				
Poss discourse	335	345	400	400
Acc discourse	330	375	395	430
Discourse effect	5	−30	5	−30

Note: *Acc* refers to accusative contexts; *poss* refers to possessive contexts.
Source: Adapted from Boland 1997b, fig. 1

decreased as the homograph became more verb dominant. When the pronoun was the possessive *his*, which allows a following noun, reading times for *and* decreased as the homograph became more noun dominant. The effects of contingent homograph bias do not appear until the word after the homograph.

Boland's research shows that frequency of use and local context of the alternative meanings of a homograph influence its immediate assignment to a syntactic category, but discourse context does not. The central question from the standpoint of LAST is, does this assignment occur as a by-product of assigning words to major phrases that correspond to thematic roles? There is no research of which we are aware that examines this question. The simplest test of this question would compare processing difficulty on the last word in sentences like the following (from Frazier and Rayner 1987; MacDonald 1993):

(41) The desert trains crashed.

(42) The desert trains soldiers.

The NVN template predicts that a sequence like *desert trains* that can be assigned the categories of NV will be assigned those categories, assuming that the noun and verb interpretations of *trains* are equally probable overall. This would make the initial assignment incorrect in the first sentence above, since *crashed* indicates that *trains* is in fact a noun, not a verb.

6.1.6 *Wh*-Trace

Wh-movement leaves behind a trace to indicate its original structural position. Here are some examples of *wh*-movement (Frazier, Clifton, and Randall 1983; Nicol and Swinney 1989):

(43) Wh-*movement*
Which book$_1$ did the teacher show t_1 to the children?

(44) The policeman saw the boy$_1$ that$_1$ the crowd at the party accused t_1 of the crime.

In the first sentence, *which book* moves from the position of object of *show* to the front of the sentence. It leaves behind a trace in the object position where it originally received the role of patient from *show*. In the second sentence, *boy* occupies the object role of *accused* in D-structure. In the form of *that*, *boy* moves to the front of the object relative clause to modify *boy* in the main clause, again leaving behind a trace that corresponds to the D-structure position of the moved element.

When *wh*-movement has occurred, there is usually an overt marker in the sentence. In the examples of *wh*-movement above, these markers are *which* and *that*. We can expect that the comprehension system may use these markers to immediately establish the trace for the moved element. The presence of overt markers for *wh*-movement distinguishes it from NP-movement, as we will see later in the section on NP-trace. However, there is another property of *wh*-movement that influences sentence processing, namely, that *wh*-movement destroys canonical structure and creates types of phrase sequences that do not correspond to any simple untransformed sentence. In chapter 5, we outlined how LAST accommodates to the unique surface phrase sequences that *wh*-movement creates. We proposed that very early *wh*-return strategies apply to activate a representation of the moved phrase in its original position. This restores a canonical surface phrase order, to which the early syntactic pattern strategies can apply. This is equivalent to using sentence pattern templates that are specific to cases of *wh*-movement.

Priming from *Wh*-trace The early reactivation of *wh*-trace correctly predicts that the *wh*-trace location primes its antecedent immediately. Swinney et al. (as reported in Nicol and Swinney 1989) used a cross-modal lexical decision task to examine the activation of *wh*-words at the location of traces (see box 6.5). Subjects heard sentences like the following and saw a test probe at a pretrace point, at the *wh*-trace, or at a posttrace point:

(45) The policeman saw the boy$_1$ that$_1$ the crowd at the party [pretrace test point] accused t_1 [*wh*-trace test point] of the [posttrace test point] crime.

The probes were either related or unrelated to *boy* (the antecedent of the trace) or to *crowd*.

Swinney et al. found significant facilitation in lexical decision for *boy* at the *wh*-trace and at the posttrace point, but not at the pretrace point (table 6.5). In contrast, there was facilitation for probes related to *crowd* at the pretrace point but not at the

Box 6.5
The cross-modal lexical decision task

The cross-modal lexical decision task involves visual presentation of a sequence of letters as the subject listens to speech. The subject's task is to indicate by voice or key press whether or not the sequence of letters spells a real word. The probe item may be related to the meaning of part of the spoken linguistic stimulus. For example,

Stimulus (speech): "The policeman saw the boy_1 $that_1$ the crowd at the party accused t_1 [T] of the crime"
Related Target (print): SON
Unrelated Target (print): HAM

One of the target words appears at a particular point in the spoken sentence—for example, at the location indicated by *T*. Typically, response times to correctly indicate that a target is a word are faster for related targets (SON) than for unrelated targets (HAM). Such an outcome is taken to show that the related word in the sentence is activated at the time of the lexical decision.

In some cases, the lexical-decision task is used with only unrelated targets to indicate processing load. In this case, a target unrelated to any word in the sentence, such as HAM for the above sentence, is presented at one of several points during the sentence. The point at which response times to make the lexical decision are longer is taken to indicate a point of increased processing load during comprehension. See Swinney 1979.

Table 6.5
Mean priming scores at three test points relative to the location of *wh*-trace (in ms, differences in lexical decision times to unrelated and related probes)

	Test point		
	Pre	*Wh*	Post
Antecedent (*boy*)	12	27	27
Control (*crowd*)	44	19	8

Source: Adapted from Nicol and Swinney 1989, table I

wh-trace or the posttrace point. The significant priming of probes related to *boy* at the *wh*-trace indicates that the *wh*-trace primes its antecedent immediately. (But see Mckoon and Ratcliff 1994 as well as McKoon, Allbritton, and Ratcliff 1996 for a different interpretation; see Nicol, Fodor, and Swinney 1994 for a reply.)

6.1.7 Subcategorization Frames

Because it appears so frequently, the NVN template is the default template. The comprehension system uses the NVN template unless there is overt and reliable contrary information. An alternative subcategorization frame of a verb can modify the

application of the NVN template, particularly if the alternative frames for the verb are used frequently.

Jennings, Randall, and Tyler (1997) demonstrated the role of frequency of use of subcategorization frames with verbs that optionally take either a noun phrase (NP) object or a sentential complement (SC):

(46) *NP object with a preferred NP verb*
The cheerful author denied the allegations.

(47) *SC with a preferred NP verb*
The cheerful author denied (that) the allegations were true.

(48) *NP object with a preferred SC verb*
The leading chemist concluded the experiments.

(49) *SC with a preferred SC verb*
The leading chemist concluded (that) the experiments were important.

Although both *deny* and *conclude* may take either type of complement, Jennings, Randall, and Tyler found that with *deny* people prefer to use a noun phrase object. With *conclude* they prefer a sentential complement.

Jennings and colleagues used the cross-modal naming task to determine whether comprehenders use subcategorization preferences online (see box 6.2). In this task, sentence fragments were presented auditorily and a target word (capitalized below) was presented visually:

(50) *Object NP target with preferred SC verb*
The leading chemist concluded . . . THEM

(51) *Subject NP target with preferred SC verb*
The leading chemist concluded . . . THEY

(52) *Object NP target with preferred NP verb*
The leading chemist denied . . . THEM

(53) *Subject NP target with preferred NP verb*
The leading chemist denied . . . THEY

The target word plausibly can be integrated with the preceding material depending on whether the main verb (*concluded* or *denied*) is interpreted as taking a noun phrase object or a sentential object. For example,

(54) The leading chemist concluded they were inert.

(55) The leading chemist denied them laboratory space.

Subjects apparently do attempt to integrate the visually presented word with the auditorily presented sentence fragment, since their naming times for targets such as

Table 6.6
Mean naming times (ms) depending on preference for use of the verb in NP object vs. sentential complement frames

	Type of target	
Preference	Object	Subject
NP object	378	388
SC	391	377

Source: Adapted from Jennings et al. 1997, fig. 2

kindly versus *kindness* were significantly faster when the target completed a grammatical sentence (421 ms):

(56) My cousin always behaves kindly.

than when it did not (434 ms):

(57) My cousin always behaves kindness.

Table 6.6 shows that naming response times were shorter numerically when the target was consistent with the subcategorization preference, but the interaction between type of target and preference was not significant. Jennings and associates did find, however, that response times to object (*them*) versus subject (*they*) targets were correlated with degree of verb preference for a sentential complement versus object noun phrase complement, as determined by pretest. For example, *indicate* shows a slight preference for object noun phrase complement, and response times to fragments with *indicate* were slightly faster for *them* than for *they*. *Accepted* strongly prefers to take an object noun phrase complement, and response times to fragments with *accepted* were much faster for *them* than for *they*. Response times were faster for *they* for verbs that prefer to take a sentential complement, and the size of this advantage for *they* increased with the strength of preference for sentential complements. Thus, comprehenders are immediately sensitive to subcategorization information, and the availability of different interpretations is related to usage preferences.

We can extend this conclusion about early sensitivity to subcategorization properties of verbs with noun phrase versus sentential objects. Similar results are found with verbs that may subcategorize for noun phrase objects versus no object. Branigan, Pickering, and Stewart (discussed in Branigan et al. 1995) report that a sentence context activates a particular subcategorization frame. They measured reading times for early-closure sentences that contained a verb that was ambiguous between transitive and intransitive (58) when another early-closure sentence preceded it (59) versus when a late-closure sentence preceded it (60). Branigan and colleagues found that reading times were faster for the early-closure target sentence when preceded by an

early-closure sentence. Similarly, late-closure targets were read faster when the context sentence was late-closure rather than early-closure.

(58) *Target sentence, early closure*
Although the film was frightening the young child enjoyed the plot.

(59) *Context sentence, early closure*
While the woman was eating the creamy soup went cold.

(60) *Context sentence, late closure*
While the woman was eating the creamy soup the pudding went cold.

This effect may involve priming of the transitive versus intransitive subcategorization frames. The early-closure context sentence primes the intransitive interpretation, which enhances the strength of a similar interpretation of *frighten* in the target sentence. This is further evidence that the comprehension system has early access to subcategorization information.

6.2 Meaning

The comprehension system uses the pseudosyntactic structure as the basis for an initial meaning. This initial meaning-form hypothesis rests on the organization of words into major phrases, the assignment of lexical categories, retrieval of likely subcategorization frames and application of canonical sequence templates. The initially accessed meaning, however, precedes the assignment of real syntax. This predicts that people are sensitive to meaning information before a complete syntactic structure.

6.2.1 Pragmatic Information Comes Before Verb Morphology

There is evidence that pragmatic violations are detected earlier than syntactic violations. Fodor et al. (1996) recorded eye movements as subjects read sentences with pragmatic violations, syntactic violations, or no violations (see box 6.6), as in the sentences below (slashes indicate the five scoring regions):

(61) *Pragmatic violation*
It seems that cats / won't usually / bake the / food we / put on the porch.

(62) *Syntactic violation*
It seems that cats / won't usually / eating the / food we / put on the porch.

(63) *No violation*
It seems that cats / won't usually / eat the / food we / put on the porch.

Table 6.7 shows first-pass residual reading times over five critical regions, depending on type of violation.

Box 6.6
The eye-tracking task

The eye-tracking task uses a camera to record the exact point on which a subject is focusing moment by moment while reading a sentence.

The eye-tracking equipment also records the amount of time that the eye rests on each point during the initial, left-to-right scan through a sentence. This time is called *first-pass fixation time* and is assumed to measure processing load at the point of initially fixating on a word.

Trueswell, Tanenhaus, and Garnsey (1994) have noted that first-pass fixation times tend to be longer for scoring regions that have more characters. However, since this correlation is not linear, it is recommended either to use scoring regions of constant length, or to correct for nonlinearity. If scoring regions are not of constant size, Trueswell and colleagues recommend correcting for nonlinearity by using *residual reading times*. Calculating residual reading times requires two steps. First, one calculates the linear regression equation that predicts first-pass reading times from length of scoring region. Second, one subtracts the predicted reading time from the actual reading time. A negative residual reading time indicates that actual reading times for a region of given length were faster than predicted, while a positive residual reading time indicates that actual reading times were slower than predicted. An alternative approach for dealing with the nonlinear relation between length of region and fixation time is to compute a quadratic function.

The eye-tracking equipment is also able to record "regressions" or backtracking. The number of regressions and amount of time spent on them are taken to indicate reanalysis and its difficulty. A third measure that the eye-tracking task yields is total processing time.

The eye-tracking procedure has some advantages and disadvantages over subject-paced word-by-word reading. The eye-tracking task does not require the subject to make a response after each word, but instead allows for more natural presentation of the reading material. However, the eye-tracking procedure does require that the subject's head be stationary, usually through the use of a chin bar on which the subject's chin rests, or a bite bar. See Carpenter and Daneman 1981; Frazier and Rayner 1982.

Table 6.7
Mean residual first-pass eye-fixation times in five scoring regions for different types of violations

	Region				
Type of Violation	1	2	3	4	5
Pragmatic	57	2	19	−20	−22
Syntactic	79	2	8	−63	−70
None	79	11	13	−52	−66

Source: From Fodor et al. 1996, fig. 1

Table 6.8
Percent of eye-movement regressions in five scoring regions for different types of violations

	Region				
Type of Violation	1	2	3	4	5
Pragmatic	5	11	12	16	29
Syntactic	7	9	19	28	20
None	8	9	6	9	15

Source: From Fodor et al. 1996, fig. 2

There was no difference in reading times in the region that contained the verb (*bake*, *eating*, or *eat*) and the following determiner (region 3). Over regions 4 and 5 (*food we* and *put on the porch*), reading times were longer for the pragmatic violations than for sentences with no violation. Reading times for sentences with syntactic violations did not differ from reading times in the sentences with no violation in any region. Thus, there was an effect of pragmatic anomaly, but not syntactic anomaly, on first-pass reading times. (Incorrect) progressive tense morphemes apparently are ignored on the first pass.

Fodor et al. found that regressive eye movements were more frequent for syntactic violations than for pragmatic violations, as shown in table 6.8. Regressions were elevated for syntactic violations in region 3 (verb + det) and in region 4 (*food we*). Finding evidence for detection of syntactic violations on regressions but not first-pass reading is problematic for syntax-first theories. Since these theories maintain that syntactic decisions are made before pragmatic decisions, we would expect the opposite, namely, effects of syntactic violations on first-pass reading, and effects of pragmatic violations on regressions. Thus, the results suggest that there is early access to pragmatic violations, and somewhat later, access to syntactic violations.

Fodor et al. also used the same materials in a cross-modal lexical decision task (see box 6.5). The same materials were presented in speech compressed to about half its normal rate. A lexical decision probe that was unrelated to the meaning of the sentence appeared at one of five points. The test points were 81 ms before the verb, at the verb, 81 ms after the verb, 162 ms after the verb, or 243 ms after the verb. Subjects indicated as quickly as possible whether or not the probe was a word. Fodor et al. interpreted slower response times on this task as evidence of increased processing load. In one condition, subjects were required to paraphrase the sentence after it was completed. Mean response times for the paraphrase group appear in table 6.9.

At the point of the verb, lexical decision times were longer for both syntactic and pragmatic violations than for no violation. In fact, at the verb and immediately after, pragmatic violations are more disruptive than syntactic violations. Only at the last

Table 6.9
Mean lexical decision times (ms) for paraphrase group at five test points depending on type of violation (test points identified by time in ms before (−) or after (+) the verb)

	Test point				
Type of Violation	−81	0	+81	+162	+243
Syntactic	794	803	770	802	810
Pragmatic	795	817	812	796	763
None	794	746	802	822	749

Source: Fodor et al. 1996, note 10

test point did lexical decision times become substantially greater for syntactic violations than for pragmatic violations.

We can understand these results in terms of LAST. The verbs *eat* and *bake* are potentially intransitive. Based on the NV pattern, the pseudosyntax presents a complete proposition, the grammar generates a candidate syntax, and the system checks it against the memory representation of the input (see figure 5.2). At the point of the verb, assigning the initial meaning draws attention away from the surface form. Hence, responses to the lexical decision task at the verb are slow. Right after the verb, the initial meaning hypothesis is available, which can then slow processing if there is a pragmatic violation. The elevated response times for syntactic violations at the end of the sentence reflect detection of a mismatch between the candidate syntax and the input. The sentences with pragmatic anomalies and no anomalies also produce syntax checking at this point, but do not produce an incorrect syntax, and hence do not result in slowed reaction times.

6.2.2 Accessing Semantic Before Syntactic Aspects of Arguments

Subcategory information provides the major syntactic category frames that a predicate requires. Some predicates require an object noun phrase, some require two object noun phrases, some an object noun phrase and a prepositional phrase, some a sentential complement, and so on. Subcategory information indicates what arguments are needed to make a complete unit of meaning. LAST therefore proposes that subcategory information and associated thematic roles are available early in order to assign major phrases to thematic roles. The question at hand is, does the semantic information become available prior to the syntactic information?

Boland (1997a) used cross-modal naming and lexical decision tasks immediately after hearing a sentence to examine the integration of a probe item with subcategorization information and argument requirements (see boxes 6.2 and 6.5). Boland used sentence fragments that contained a verb that was ambiguous between taking two or three arguments (*toss*). Other sentence fragments contained a verb that unambig-

uously takes two (*inspect*) or three arguments (*describe*). (Some readers may feel *describe* can take only two arguments, as in *Nancy described the necklace.*) Sample materials appear below, with the probe item in capital letters:

(64) *Ambiguous, two-argument*
Which salad did Jenny toss BILL

(65) *Ambiguous, three-argument*
Which baseball did Jenny toss BILL

(66) *Unambiguous, two-argument*
Which necklace did Nancy inspect SAM

(67) *Unambiguous, three-argument*
Which necklace did Nancy describe SAM

The subcategorization and thematic requirements of these verbs allowed Boland to examine the availability of semantic and syntactic information. Semantically, the verb *toss* may take either two arguments (agent, patient) or three arguments (agent, recipient, and patient):

(68) Jenny tossed the salad.

(69) Jenny tossed Bill the ball.

Syntactically, the subcategorization frame of *toss* is ambiguous between

toss ⟨NP⟩

toss ⟨NP NP⟩

where the material inside the angle brackets indicates phrases that are required inside the verb phrase (i.e., omitting the subject, which is always required).

In Boland's (1997a) analysis, *inspect* semantically takes exactly two arguments (agent, patient), and *describe* takes three (agent, patient, and goal):

(70) Nancy inspected the necklace.

(71) Nancy described the necklace to Sam.

Syntactically, the subcategorization frames of these verbs are:

inspect ⟨NP⟩

describe ⟨NP PP⟩

Since the subcategorization frame of *inspect* does not allow it to take a second argument inside the verb phrase, we expect that it will be difficult to integrate the probe *Sam* into the sentence. However, the subcategorization frame of *describe* does allow it to take an additional argument, but, syntactically, it must be in the form of a prepositional phrase, not a noun phrase, since neither of the following is grammatical:

Table 6.10
Mean naming times (ms) for ambiguous and unambiguous verbs depending on the number of arguments they allow

	Arguments		
	2	3	Mean
Ambiguous	609	608	609
Unambiguous	627	625	626

Source: Adapted from Boland 1997a, table 2

(72) *Nancy described the necklace Sam.

(73) *Nancy described Sam the necklace.

Thus, *describe* can take a goal as well as an agent and patient, and, on semantic grounds, Sam is a plausible goal, but the third argument must be in the form of a prepositional phrase, not a noun phrase.

Boland used the naming task as a measure of sensitivity to syntactic constraints and the lexical decision task as a measure of sensitivity to semantic constraints (see Boland 1993). Because of the differences in subcategorization and thematic requirements, sentences with different verbs will produce differences in performance on these tasks. Consider first the naming task, which other studies have shown to be sensitive to syntactic constraints (see box 6.2). If all subcategorization frames are momentarily and indifferently accessed, there will be no difference in naming times for the two ambiguous sentences with *toss*. Since neither *inspect* nor *describe* subcategorizes for a third NP, there also will be no difference in naming times for *inspect* and *describe*. Since the syntactic requirements of *inspect* and *describe* do not include a third NP, naming times for these sentences will be slower than naming times for the ambiguous *toss*, which does subcategorize for a third NP. The results for the naming task appear in table 6.10.

Boland found that naming times were faster for ambiguous sentences than for unambiguous sentences. This shows that both subcategorization frames of the ambiguous verb are available for integration with the probe. Boland also found that response times did not differ between the ambiguous two-argument condition (*toss* with *salad*) and the ambiguous three-argument condition (*toss* with *baseball*). If plausibility influenced the availability of subcategorization frames, one might expect that *salad* would rule out the three-argument frame. Because response times did not differ for the two- and three-argument ambiguous conditions, Boland concluded that all syntactic subcategorization frames possible for a verb are activated, regardless of semantic context. However, we could as easily conclude that the syntactic frames are

Table 6.11
Mean lexical decision times (ms) for ambiguous and unambiguous verbs depending on the number of arguments they allow

	Arguments	
	2	3
Ambiguous	659	643
Unambiguous	699	666
Mean	679	655

Source: Adapted from Boland 1997a, table 3

not accessed at all in a way that influences the naming time—that is, the differential effects of syntactic complexity do not have an immediate effect.

Boland's evidence unequivocally shows that the semantic aspects of the argument structure are immediately accessed. Consider the lexical decision task, which is sensitive to semantic constraints (see box 6.5). If meaning is accessed initially, lexical decision times will be faster for the three-argument verbs than for the two-argument verbs, since in the three-argument cases the probe can be interpreted with the sentence fragment to complete the argument requirements of the verb. Thus, response times ought to be faster for the *toss* sentence with *baseball* and the *describe* sentence than for the *toss* sentence with *salad* and the *inspect* sentence.

The results for the lexical decision task appear in table 6.11. Boland found that lexical decision times were faster for three-argument sentences than for two-argument sentences. This result suggests that the probe item is integrated semantically with the three-argument sentences, which is not possible with the two-argument sentences. Based on the argument requirements of subcategorization frames, the probe item is incorporated whenever possible into an initial representation of meaning, regardless of the detailed syntactic requirements of the verb.

6.3 Real Syntax

Up to this point, we have highlighted experimental facts consistent with an early formulation of a meaning-form hypothesis, prior to a complete syntax. We now turn to evidence for the regeneration of a complete syntactic description from constraints set by the formulation of the initial meaning. The real syntax of a sentence refers to the entire derivation and annotated phrase structure tree for the sentence. The output of real syntax differs from pseudosyntax in two ways. First, real syntax includes detailed phrase structure within major phrases, whereas pseudosyntax includes the segregation of words into major phrases based on function words and major lexical categories. Second, real syntax includes the linking of NP-traces and antecedents.

The grammar generates a real syntax for a sentence based on an initial guess about meaning. This initial meaning-form hypothesis includes the thematic roles that have been associated with major phrases.

6.3.1 Well-Formedness Constraints Apply After Meaning and Pseudosyntax

Semantic information appears early enough to influence performance on a sentence matching task (see box 6.7). As subjects performed the sentence matching task, Murray and Rowan (1998) measured first-pass fixation times in an eye-tracking task (see box 6.6). The matching sentences were designed so that the initial noun phrase/verb pair was either plausible or implausible, and so that the verb/second noun phrase pair was either plausible or implausible. Examples of their materials appear below:

(74) *Plausible NP_1-V, plausible V-NP_2*
The hunters stacked the bricks.

(75) *Plausible NP_1-V, implausible V-NP_2*
The hunters stacked the tulips.

(76) *Implausible NP_1-V, plausible V-NP_2*
The bishops stacked the bricks.

(77) *Implausible NP_1-V, implausible V-NP_2*
The bishops stacked the tulips.

The basic measure of interest in this version of the sentence-matching task was first-pass reading time on the initial noun phrase in the first sentence. The results are most interesting, and appear in table 6.12. For the initial noun phrase in the first sentence, first-pass reading times were 15 ms/word faster for plausible NP_1-V sequences than

Box 6.7
The sentence-matching task

In the sentence-matching task, the subject sees one whole sentence, then presses a key to display a second sentence so that both sentences appear simultaneously. In some experiments, the second sentence appears after a certain time, such as 2 sec, without the subject pressing a key. The two sentences appear together centered and on adjacent lines, as in

The hunters stacked the bricks.
The hunters stacked the bricks.

The subject's task is to indicate whether the two sentences contain the same words in the same order. The dependent variable typically is response time to make this judgment, measured from the onset of the second sentence. See Forster and Olbrei 1973.

Table 6.12
Mean first-pass reading times (ms/word) on the initial noun phrase (NP_1) in the first sentence

	V-NP_2		
NP_1-V	Plausible	Implausible	Mean
Plausible	244	238	241
Implausible	252	260	256
Difference	8	22	15

Source: Adapted from Murray and Rowan 1998, table II

for implausible NP_1-V sequences. This difference was significant despite the fact that subjects were not required to understand the sentences, only to match them physically. The results suggest that semantic information becomes available quickly and mandatorily, indicating that meaning is an early part of comprehension. Murray and Rowan argue that the semantic-plausibility effect is due to computation of the semantic relation between the initial noun phrase and the verb, which does not occur early according to LAST. The effect, however, may have occurred because of immediate access to associative information, such as residual associations between the noun and the verb or the plausibility of the initial noun as an agent, both of which are elicited early according to LAST. If *hunters*, for example, is a more plausible agent than *bishops*, a sequence that begins with *hunters* conforms more strongly to the initial noun = agent template (see also section 7.1.1). Thus, we would expect faster processing time on *hunters* than on *bishops*, regardless of the verb. Supporting this interpretation was a significant semantic effect on fixation times for the initial noun phrase *before* the verb had been fixated.

Freedman and Forster (1985) showed that some aspects of syntax are accessed in the sentence-matching task. These aspects of syntax include the organization of words into major phrases that correspond to arguments, subject-verb agreement, and quantifier placement—that is, those features marked by pseudosyntax. Their materials included acceptable and unacceptable sentences, and their unacceptable sentences involved violations of phrase structure, agreement, and quantifier placement rules. The phrase structure violations consisted of either two noun phrases or two verb phrases from two different acceptable sentences. Some sample materials are:

(78) *Acceptable phrase structures*
   a. The girl behind you reminds me of your sister.
   b. The subsequent discussion soon got boring.

(79) *Phrase structure violations*
   a. The girl behind you the subsequent discussion.
   b. Reminds me of your sister soon got boring.

Table 6.13
Mean sentence-matching times (ms) for three violations of pseudosyntax

	Phrase structure	Agreement	Quantifier
Acceptable	1432	1385	1274
Violations	1555	1427	1340
Difference	123	42	66

Source: Adapted from Freedman and Forster 1985, tables 5 and 6

(80) *Acceptable agreement*
Mary was writing a letter to her husband.

(81) *Agreement violation*
Mary were writing a letter to her husband.

(82) *Acceptable quantifier placement*
The baby ate his cereal all up.

(83) *Quantifier placement violation*
The baby ate his cereal up all.

As in the sentence-matching task (box 6.7), Freedman and Forster presented two lines of text on a screen. In the critical positive trials in which the lines matched, the lines were either acceptable or unacceptable sentences. Freedmen and Forster recorded response times to judge whether the two lines were identical. No judgment of acceptability was required. The results appear in the table 6.13. For each type of material, response times were significantly longer for violations than for the corresponding acceptable sentences. These results suggest that constraints on phrase structure, subject-verb agreement, and quantifier placement become available quickly and mandatorily.

While meaning and pseudosyntax are available early, certain syntactic information becomes available only at a late stage of comprehension. This late syntactic information includes well-formedness constraints that block the grammar from allowing certain kinds of movement. Consider the examples below from Freedman and Forster 1985:

(84) *No movement, no specified subject*
The duchess sold a portrait of her father.

(85) *Movement, no specified subject*
Who$_1$ did the duchess sell a portrait of t_1?

The examples above show that English syntax allows movement of a *wh*-word that refers to *father*. In these examples, the movement crosses a nonspecific subject, *a portrait*. However, the corresponding movement across a specific subject, *Turner's*

portrait in the examples below, produces an ungrammatical sentence:

(86) *No movement, specified subject*
The duchess sold Turner's portrait of her father.

(87) *Movement, specified subject*
*Who$_1$ did the duchess sell Turner's portrait of t_1?

Since the grammar must allow movement of the object in some cases but not others, there must be a filter that rules out certain cases. The relevant difference between the grammatical and ungrammatical cases is that sentences in which the movement crosses a specific subject such as *Turner's portrait* are ungrammatical. Crossing a nonspecific subject (*a portrait*) is allowed. This constraint is called the *specified-subject constraint*.

Another example of a violation of a well-formedness constraint is subjacency. The grammar allows movement across one S-node, indicated below with left bracket, since there are acceptable sentences such as the following (Freedman and Forster 1985):

(88) *No movement, no subjacency constraint*
It is hard to trust Ann.

(89) *Movement, no subjacency constraint*
Ann$_1$ is hard [to trust t_1].

But moving an element across two NP- or S-nodes produces an ungrammatical sentence:

(90) *No movement, subjacency constraint*
It is hard to trust rumors about Ann.

(91) *Movement, subjacency constraint*
*Ann$_1$ is hard [to trust [rumors about t_1]].

The subjacency constraint therefore prohibits sentences in which there has been movement of an element across two NP- or S-nodes.

Freedman and Forster (1985) used the sentence-matching task to examine sensitivity to violations of the specified subject and subjacency constraints (see box 6.7). The distractor trials were pairs of sentences of similar construction in which the sentences differed in just one word.

In order to assess sensitivity to well-formedness constraints on movement, it was necessary to compare the effects of movement with violations of a constraint and the effects of movement without violation of a constraint. If subjects are sensitive to these constraints in the sentence-matching task, there should be greater effects of movement on matching times for sentences with violation of a constraint than for sentences without violation a constraint. The results for critical "same" responses

Table 6.14
Mean sentence-matching times (ms) for specified subject constraint and subjacency

	No violation	Violation	Difference
Specified subject			
Movement	1268	1549*	281
No movement	1199	1511	312
Difference	69	38	
Subjacency			
Movement	1198	1498*	300
No movement	1050	1350	350
Difference	148	148	

Note: Asterisks indicate the ungrammatical cases.
Source: Adapted from Freedman and Forster 1985, tables 1 and 3

appear in table 6.14. For both specified subject and subjacency constraints, there was no difference in the effect of movement on response times whether a constraint was violated or not. These results suggested that subjects can perform the sentence-matching task without accessing knowledge about the specified subject and subjacency constraints.

However, subjects do access these constraints late in comprehension. An unspeeded test—that is, allowing subjects to take their time—showed that subjects judged sentences with movement that violated the specified subject constraint to be odd or bad in 82 percent of the cases, compared to 33 percent for sentences with movement that did not violate this constraint. Thus, the specified subject constraint is accessed late, after the response is made on a speeded sentence-matching task.

The results of this set of studies show that meaning and pseudosyntax appear early enough to influence judgments on the sentence-matching task, but constraints on movement do not. The sentence-matching task requires comparison of two sentences based *only* on lexical items. Nevertheless, certain higher-level properties of sentences influence matching performance. These properties include the organization of words into major phrases such as noun phrases and verbs and the local plausibility of the subject phrase and the verb. What does not influence matching are high-level constraints on movement, even though they do influence unspeeded judgements of grammaticality. The sentence matching results provide clear evidence that certain aspects of syntax are accessed late.

6.3.2 Early Access to Major Phrase Segmentation, Late Access to Complete Phrase Structure

A series of studies on the mislocation of brief sounds (clicks) presented during sentences supports the view that major phrases are isolated early but detailed phrase

Box 6.8
Response bias in the click-location task

In early click-location tasks, the subjects listened through headphones to a spoken sentence in one ear and a click in the other ear. They then wrote down the sentence and marked where the click had occurred. Errors in locating the click were attributed to processes that occur during the perception of the sentence. An alternative "response bias" interpretation is that errors in click location occur because of the subject's use of phrase structure as a cue to where misperceived clicks must have occurred.

Bever, Hurtig, and Handel (1975) addressed the problem of response biases directly. In their procedure, subjects listened to the sentence and click. They then turned the page of a booklet to reveal the sentence already printed, with a "window" of location for the click, as marked by slashes below.

Along with /his wives the prince/ brought the court's only dwarf.

The subject's task was to mark the subjective location of the click within that window. The window was always centered on the objective click location, which was either one word before the clause break, in the clause break, or one word after the clause break. However, in some trials with experimental sentences, the click was absent. Bever, Hurtig, and Handel prepared subjects for this by including nonexperimental sentences with clicks of varying loudness. This gave the subjects the information that the clicks varied in loudness and sometimes were hard to hear. In addition, to keep subjects vigilant about encoding the entire sentence, the page revealed for occasional nonexperimental sentences was blank, and subjects had to write down the entire sentence and locate the click in it.

This procedure allows direct comparison of the pattern of mislocations when the click was present versus absent. Bever and associates found significantly different patterns in the click-present and click-absent trials. In the click-absent trials, the position just before the clause break attracted the most responses. In the click-present trials, many more responses were located in the phrase break than either preceding or following it. Thus, the mislocation pattern to actual clicks is perceptually mediated. See Bever 1973; Garrett, Bever, and Fodor 1966.

structure is assigned late. The paradigm is described in box 6.8. Fodor and Bever (1965) reported that clicks are mislocated towards clause boundaries. The characteristic response pattern was that clicks objectively prior to or after the clause-break position [O] were reported as having occurred in the clause break [R]:

Along with his wives the prince brought the court's only dwarf.
O R

Correspondingly, clicks objectively in the clause-break position were accurately located more often than clicks in other positions.

Along with his wives the prince brought the court's only dwarf.
OR

Fodor and Bever initially proposed that "surface phrase structure" was assigned immediately during comprehension and that phrase units tended to resist interruption, thus accounting for the displacement of clicks to points between clauses. They argued that their data justified the claim that the entire phrase structure is assigned immediately, because assigning the clause break logically occurs as a function of that, but only if the complete structure is assigned. In LAST, the assignment of clause segmentation occurs as a consequence of assigning basic thematic relations to arguments of verbs. A complete surface phrase structure is not necessary for that.

Bever, Lackner, and Kirk (1969) examined the question of whether surface phrase structure segmentation or underlying thematic roles affect click location. They contrasted sequences that differed in their "deep" structure but not in their surface structure:

(92) The corrupt police can't force criminals [to confess].

(93) The corrupt police can't bear [criminals to confess].

The phrase structure relation between the main verb (*bear*, *force*) and the following noun phrase (*criminals*) is the same in these two sentences. However, in the first sentence the noun phrase is the patient of the main verb and the agent of the complement verb. In the second sentence the entire complement is the patient of the main verb, while the noun phrase is still the agent of the complement verb. Thus, the only difference is whether or not the noun phrase is isolated as the patient of the verb. Bever, Lackner, and Kirk found that clicks objectively located in the main verb were more often misplaced into the position between the verb and the noun phrase in the second sentence (80%) than in the first sentence (40%). This result is consistent with the view that the argument structure and associated segmentation can govern click mislocation.

Bever, Lackner, and Kirk also examined whether all structural details are initially assigned or just the phrases that correspond to the complete set of thematic roles. In one experiment, they used materials like:

(94) When he stood up my son's book fell off the table.

Clicks were objectively located within one of the two words preceding the clause boundary (*stood*, *up*) or within one of the two words following the clause boundary (*my*, *son's*). Bever and colleagues classified the mislocations in terms of whether or not they were toward larger breaks in the phrase structure tree. The relevant structural details here are that *stood up* constitutes a phrase structure unit, but *he stood* does not. Thus, if detailed phrase structure is assigned immediately, clicks within *stood* will be mislocated more often before *stood* than after it.

Bever, Lackner, and Kirk found that less than 42 percent of the mislocations supported the within-clause phrase structure hypothesis. Thus, while clicks objectively in

words tend to be reported as having occurred between words in general, there was no effect of minor phrase boundary differences.

These results created a conundrum:

1. Clicks within words are subjectively reported between words.
2. Only major phrases that correspond to the required arguments of a verb affect click location.

That is, words are units relevant for click mislocation, and major phrases are also relevant, but minor phrase structure that rests between the level of words and major phrases is not relevant. This conundrum either invalidates the idea that the entire surface phrase structure determines subjective click location, or it invalidates syntax-first theories. However, it is entirely consistent with LAST. The system simultaneously assigns arguments to verbs, which requires at least isolating major phrase heads. At the same time, to map an initial meaning-form hypothesis, it must isolate as many content words as possible. Thus, the system accesses words and major phrases as an initial part of its computation.

The initial assignment of major argument relations requires that they be identifiable in the surface sequence. This process of idenifying major argument relations involves an initial process of segmentation of lexical sequences into likely major phrases, based on superficial structural cues such as the location of function words. This in turn depends on segregating and representing separately function words and content words. These early steps in comprehension are supported by the systematic mislocation of clicks while listening to sentences. Clicks are systematically mislocated as occurring between words, and also between major argument-bearing phrases.

These results would also be consistent with a full phrase structure–building component, of the sort presupposed by most syntax-based comprehension theories. But clicks are not perceptually attracted to phrase structure breaks intermediate between word- and argument-phrase level. The failure initially to assign a complete structure is consistent with the LAST. But it is also a prediction of the LAST that the complete structure *is* assigned late, as part of the grammatical synthesis of the sentence for comparison with the input. Thus, the prediction is that minor variations in phrase structure will affect click mislocation, but only after some time, not as an immediate perceptual effect.

Bever and colleagues tested this using a standard click-location paradigm but requiring subjects to wait a few seconds after hearing the sentence and auditory click before locating it. With this paradigm, minor variations in phrase structure do have the predicted systematic effect on click mislocations, suggesting that the entire phrase structure is now mentally represented. It might be argued that all the effects of structure on click location *must* be a response bias, since locating a click clearly necessitates postperceptual process. This is not logically required and should be subjected to the same kind of test as the mislocation of clicks to major boundaries

(Bever, Hurtig, and Handel 1975). If future research shows that nonexistent clicks are "located" at minor phrase boundaries a few seconds after hearing the sentence, this will demonstrate that some kind of postperceptual process is at issue. Of course, LAST claims that this postperceptual process is specifically the reconstitution of the entire syntactic derivation including the surface phrase structure, so even if it were a response bias the systematic pattern is consistent with the model.

6.3.3 NP-Trace

The properties of NP-movement provide further evidence that the details of syntax are assigned after meaning. We noted in chapter 4 that NP-movement differs from *wh*-movement of major arguments. NP-movement does not leave explicit markers of movement, and it preserves canonical order. This allows NP-trace sentences to be understood initially by (mis)application of the initial template. These two properties of NP-movement suggest that the comprehension system establishes NP-traces in their original location *after* assigning a meaning, rather than before as it does in the case of *wh*-movement (see section 6.1.6).

Here are some examples of NP-movement (from McElree and Bever 1989):

(95) *NP-movement*
   a. The judge$_1$ is certain t$_1$ to argue the appeal.
   b. The lawyer$_1$ was suspected t$_1$ by the judge.

In the first sentence above, *judge* occupies the subject position in front of the verb *argue*, where it receives its thematic role as agent of *argue*, When *judge* is moved to its surface position at the beginning of the sentence, it receives nominative case assignment. Because of the grammatical requirement in government and binding theory that thematic roles are projected at all levels of structure, the movement operation leaves behind a trace t_1. The common subscript indicates that *t* and *judge* refer to each other.

In the second sentence, *lawyer* occupies the object position following *suspected* in D-structure, where it receives the role of patient of *suspected*. There is movement to the front of the sentence, leaving behind a trace following the verb. Once again the common index indicates coreference.

NP-movement creates sequences similar to base-generated sequences. Consider the following pair of sentences:

(96) *NP-movement*
   a. The judge$_1$ is certain t$_1$ to argue the appeal.

   PRO
   b. The judge is eager PRO to argue the appeal.

The first sentence above has NP-movement. However, the sequence of major phrases in the first sentence corresponds to those of the base-generated sequence in the second

sentence. The deep subject of *is eager* is *judge*, but the deep subject of *is certain* is *the judge to argue the appeal*. The following pattern of acceptable and unacceptable sentences confirms that this pair of sentences has different main subjects:

(97) a. It is certain that the judge will argue the appeal.
b. *It is eager that the judge will argue the appeal.
c. For the judge to argue the appeal is certain.
d. *For the judge to argue the appeal is eager.

Similar differences occur in the following pair, with the first sentence but not the second having NP-movement:

(98) a. The lawyer$_1$ was suspected t_1 by the judge.
b. The lawyer was unsatisfied by the decision.

In the first sentence above, *lawyer* is the deep object of *suspected*, but in the second sentence it is not the deep object of *unsatisfied* because *unsatisfy* is not a transitive verb. The following pair confirms that *lawyer* is the deep object only in the first sentence above:

(99) a. The judge suspected the lawyer.
b. *The decision unsatisfied the lawyer.

In the case of sentences with NP-movement the comprehension system may initially assign an incorrect syntactic structure while accessing an almost correct meaning. The similarity of surface patterns between sentences with NP-movement and other base-generated patterns may lead to an initial meaning-form hypothesis that is based on canonical sentence patterns that have different meanings. For example, the following sentence pattern may elicit an initial meaning-form hypothesis for predicate adjective sentences:

(100) NP + BE + ADJ $\Rightarrow$ Experiencer + BE + state

Applied to the sentence

(101) The lawyer was unhappy with the decision.

the template yields the following meaning assignments:

lawyer = Experiencer
unhappy with the decision = State

Applied to a passive sentence, the same template yields corresponding meaning assignments:

(102) The lawyer was suspected by the judge.

lawyer = Experiencer
suspected by the judge = State

Because of the similarity in surface patterns between passive and predicate adjective sentences, a passive sentence is interpreted initially as a stative (adjectival) sentence like *the lawyer was unhappy*. This hypothesis about the meaning serves as the basis and constraint on what the grammar generates as a corresponding syntactic derivation. The grammar attempts to generate the syntax for a stative based on grammatical information associated with the specific lexical items that have been recognized. However, the grammatical information associated with *suspect* includes the fact that it is a transitive verb, which requires an object (patient) and, optionally, a subject (agent). Thus, the grammar will be unable to generate an adjectival structure, and will block the analysis, disposing of the initial meaning-form hypothesis. The comprehension system receives the information that the grammar failed to generate an adjectival structure, and accessing the subcategorization frames allowed by the verb derives the hypothesis that the meaning is

judge = agent
suspect = action
lawyer = patient

Based on this meaning, the grammar generates the correct syntax.

Thus, sentences with NP-movement will be processed differently than sentences with *wh*-movement. Special strategies for developing a conceptual representation for *wh*-sentences lead to generation of the correct syntax. The similarity of sentences with NP-movement to other canonical sentences leads to generation of an initially incorrect syntax. Since the initial hypothesis about meaning for sentences with NP-movement leads to the incorrect syntax, LAST assigns the correct syntax of such sentences late.

Priming from NP-Trace Experimental evidence supports the prediction that the correct syntax for sentences with NP-movement is assigned late. McElree and Bever (1989) used a probe recognition technique with sentences like those below (see box 6.9):

(103) *NP-movement*
The dazed cabbie$_1$ who drove the beat-up taxi was resented t_1 (P1) constantly (P2).

(104) *Predicate adjective*
The dazed cabbie who drove the beat-up taxi was resentful (P1) constantly (P2).

P1 and P2 indicate the points at which the probe items appeared. The first sentence has a passive structure, with a trace indicating the deep position of *the dazed cabbie*. The second sentence has a predicate adjective structure, which has no movement and

Box 6.9
The probe-recognition task

In the probe-recognition task, the subject hears or reads linguistic material. At some point the material stops and the subject hears or sees a probe item. The subject's task is to indicate as quickly as possible whether or not the probe had appeared in the sentence. Timing usually begins with the onset of the probe item, and ends with the subject's vocal or manual response. For example, subjects might hear

Though Pete called up his aunt each night at [P] nine, he rarely called his grandmother.

and see a probe such as UP at the point corresponding to P (Townsend and Bever 1978). By presenting the probe task at different test points during the sentence, it is possible to determine accessibility to the target item at different points. Rapid response times are taken to indicate that the target is easily accessed at that test point. See also Caplan 1972.

Table 6.15
Mean probe recognition times (ms) in passive vs. adjectival sentences at two test points

	Test point relative to NP-trace	
Sentence	Trace	Posttrace
Passive	901	932
Adjectival	909	1008

Source: From McElree and Bever 1989, table II, exp. 5

no trace. While reading such sentences, subjects received a probe word such as DAZED at the point of P1 or P2. The results appear in table 6.15.

McElree and Bever (1989) found no difference in response times at the first probe position for adjectival and passive sentences. At the second probe position, however, response times were faster for passive sentences than for adjectival sentences. These results indicate that the correct syntax of a passive sentence, including the trace, is not assigned immediately. The initial interpretation of a passive sentence corresponds to that of a predicate adjective sentence.

In a separate experiment, McElree and Bever (1989) tested for antecedent activation in NP-raising sentences. Example sentences are:

(105) *NP-raising*
The conceited actor$_1$ who worked with the leading lady was sure t_1 to (P1) rehearse for the entire evening (P2).

(106) *PRO*
The conceited actor who worked with the leading lady was eager PRO to (P1) rehearse for the entire evening (P2).

Table 6.16
Mean probe recognition times (ms) in three sentence types at two test points

	Test point	
Sentence	P1	P2
NP raising	882	922
PRO	909	986
Adjectival	905	1022

Source: From McElree and Bever 1989, table II, exp. 2

(107) *Adjectival*
The conceited actor who worked with the leading lady was rude to (P1) the rehearsers in the evening (P2).

The results appear in table 6.16. A NP-trace had no immediate effect on probe recognition response times, but it did have an effect at the end of the sentence. Response times for the three sentence types did not differ at the NP-trace position, but they were faster at a later point for NP-raising sentences than for either PRO or adjectival sentences.

The results show that NP-trace does not activate its antecedent at the trace position, but it does so considerably later. This suggests that the initial interpretation of NP-raising is similar to a predicate adjective sentence. Later, the comprehension system has assigned the correct syntax of *the conceited actor* as subject of *rehearse*.

Osterhout and Swinney (1993) confirmed the late effect of a NP-trace in passive sentences with a cross-modal lexical decision task. They presented a single item for lexical decision during a spoken sentence such as

(108) *Passive*
The dentist$_1$ from the new medical center in town was invited t_1 by the actress to go to the party.

(109) *Active*
The dentist from the new medical center in town invited the actress to go to the party.

A lexical decision probe related to the meaning of *dentist*, such as *tooth*, occurred immediately at the point of the trace (t_1), 500 ms later (in the passive sentence, near the beginning of *actress*), or 1000 ms later (in the passive, near the end of *actress*). Unrelated probe words such as *flood* appeared at the same points. The results appear in table 6.17.

In passive sentences, only at 1000 ms did related probes produce significantly faster response times than unrelated probes. In active sentences there was no significant

Table 6.17
Mean lexical decision times (ms) in passive and active sentences at three test points

	Test point relative to NP trace		
	0 ms	500 ms	1000 ms
Passive			
Related	589	637	615
Unrelated	607	661	661
Difference	18	24	46
Active			
Related	592	657	645
Unrelated	600	654	653
Difference	8	−3	8

Source: Adapted from Osterhout and Swinney 1993, table III

difference between related and unrelated probes at any of the three test points. These results support LAST. NP-trace primes its antecedent, though this priming does not occur immediately. The mechanism for priming occurs after the formation of an initial guess about meaning and after generation of the incorrect syntax, at the stage of revising the syntax.

Word-priming results like these could, of course, be due to factors other than the syntax. For example, the passive construction in English characteristically gives discourse focus to the surface subject, deep patient. In that view, it is not surprising that there is selective priming to the subject. Of course, this does not explain the lack of priming for noneventive passives, which also focus the surface subject and it does not explain why NP-traces are primed late. But it is useful to explore NP-trace priming with other constructions that might control for some of the potential artefacts. Bever and Sanz (1997) did this in Spanish. Sanz (1996) noted that unaccusative intransitive verbs in Spanish are syntactically raising verbs (verbs such as *fall*, *arrive*). This follows from the fact that such verbs are telic, and must be delimited by a verb phrase–internal argument, but unergative intransitive verbs do not exhibit evidence of raising (verbs such as *speak*, *sleep*). Thus, Bever and Sanz (1997) could contrast priming for the initial noun phrase in sentences like the following:

(110) *Unaccusative*
El apuesto critico … llego [np]. …

(111) *Unergative*
El apuesto critico … hablo. …

In critical cases, Bever and Sanz (1997) found more priming for unaccusative than unergative verb constructions. This result is important because the initial noun phrase position has the same discourse role in Spanish for both types of verbs.

6.3.4 Synthesis Based on Conceptual Representation

LAST relies on a generative component to synthesize a detailed syntactic representation. The input to the synthetic component—that is, the grammar—has available a representation of meaning, words, and bits of a structural representation. This component of the model resembles the computation of syntactic representations in sentence production, in which the speaker begins with an intention to utter some conceptual representation and must then formulate the structure of a sentence that expresses it. However, unlike sentence production, the representation that the syntactic synthesis supplies for comparison with the input is itself internal, not an actual utterance.

The synthetic component of comprehension provides an internal syntactic representation of a sentence, based on a conceptual representation and the actual words. When we memorize a sentence for later recall, the possibility arises that it is encoded abstractly and is fleshed out into an actual utterance only when actively recalled. The classic studies of sentence memory by Miller and his students (chapter 2) suggested that sentences are coded in terms of basic propositions. Consistent with this view, Lombardi and Potter (1992; Potter and Lombardi 1990, 1998) proposed the *regeneration hypothesis* for sentence recall: When a sentence is recalled verbatim, the recall is based on a conceptual representation, including an abstract representation of the verb and its argumetns, and activated lexical items. The recalled sentence is generated from this conceptual representation using the rules of grammar.

Lombardi and Potter (1992) used a RSVP recall task following a list of distractor words (see box 6.10). The duration of the words was 200 ms, and the sentence was followed by a sequence of five distractor words. In the critical cases, the list of distractors contained a word related to a target word in the sentence. Their procedure involved presenting a sentence in RSVP (200 ms/word) followed by a sequence of five words, and then a probe item. The subject must indicate whether the probe item was in the sequence of five words, and then recall the sentence as accurately as possible. The critical sentences contained alternating datives, which can present the recipient as a prepositional phrase or as a noun phrase:

(112) *NP-PP*
The rich widow is going to give a million dollars to the university.

(113) *NP-NP*
The rich widow is going to give the university a million dollars.

Box 6.10
The RSVP recall task

In the Rapid Serial Visual Presentation (RSVP) task each word of a sentence is displayed for a brief amount of time, one after the other. The word-presentation times typically are constant within a sentence, and they range from durations as short as 20 ms in studies of unconscious word priming to durations as long as 250 ms. Participants typically are asked to recall immediately as many words as possible. In some uses of this paradigm, the sentence is followed by a sequence of distractor words, and then the instruction to recall the sentence appears. The distractor words prevent the participants from recalling the last words of the sentence out of a short-term memory buffer. See Forster 1970.

The sequence of five words contained a lure such as *donate*. The lure is a nonalternating dative that can appear only with the recipient in a prepositional phrase. Lombardi and Potter assume that if the surface structure is stored, the lure will intrude more in sentence recall when it matches the surface frame of the sentence to be recalled. If a conceptual representation is stored, lure intrusions will not differ for the two types of sentences.

Lombardi and Potter (1992) found no significant difference in intrusions of the lure in the two types of sentences. For NP-PP sentences the intrusion rate was 11 percent, and for NP-NP sentences it was 7 percent. Furthermore, when the lure did intrude in recall of NP-NP sentences, the grammatical form of the sentence was changed to fit the grammatical requirements of the lure. For NP-PP sentences, changes in grammatical form occurred in only 7 percent of the cases.

In a second experiment, subjects read a sentence in RSVP, then read a probe word, and had to say whether the probe could replace a word in the sentence without changing anything else in the sentence, including the meaning. When the subjects said yes, they then had to report the sentence as it would appear with the probe word. The sentences were alternating datives in either NP-PP or NP-NP form:

(114) *NP-PP*
The agent will send his report to the government when his mission ends.

(115) *NP-NP*
The agent will send the government his report when his mission ends.

The probe words were either alternating datives (*mail*) or nonalternating datives (*transmit*).

The results appear in table 6.18.

When the probe word was an alternating dative, the probe could fit into either type of sentence. In this case, accuracy was high and response time fast. When the probe word was a nonalternating dative, accuracy was high for NP-PP sentences but not for

Table 6.18
Percentage of correct responses and mean RTs (ms)

	% Correct	RT
Alternating probe		
NP-PP sentence	88	1460
NP-NP sentence	88	1434
Nonalternating probe		
NP-PP sentence	91	1379
NP-NP sentence	33	1829

Source: Adapted from Lombardi and Potter 1992

NP-NP sentences. The correct response for the NP-NP sentences is no, since substituting the probe into the sentence would yield an ungrammatical sentence:

(116) *The agent will transmit the government his report when his mission ends.

Since accuracy was so low for trials with NP-NP sentences and nonalternating probes, the subjects must not have stored the surface structure. The percentage of actual changes in the sentence confirms this conclusion. When subjects incorrectly said yes, their recall of the sentence with the probe verb involved a change to restore grammaticality in four times as many cases as when they said no.

While almost tautological, the Lombardi and Potter 1992 study gives some empirical force to an existence proof for corresponding processes in the synthetic component of comprehension. To put it in terms relevant to the work of Potter and Lombardi: If sentence comprehension involves mechanisms of literal recall, the recall process involves a synthesis of the surface form from conceptual representations together with specific content words, verbs, and their arguments. Garrett (1999) has generalized this notion, arguing that many component processes of speech production are involved in assigning structure during comprehension.

6.4 Introspection and Sentence Processing

Linguistic theory relies largely on intuitions about well-formedness of sentences, independent of their processing complexity. Accessing an acceptability intuition calls on the native speaker's direct knowledge of whether a sentence initially appears to be computable as a sentence. While such intuitions can be refined by experience or theoretical expectations and are sometimes more variable than one would like in data, there is a bedrock of stable intuitions underlying syntactic theory. There is a sense in which grammaticality intuitions involve examination of the language, not of the speaker's attitude toward or behavior with the language. This contrasts with intu-

itions about relative complexity, which are sometimes used to initiate or buttress processing theories. To render an intuition about comprehension complexity, a speaker must imagine perceiving the sentence (or be presented with it for comprehension), and then introspect about how difficult it seemed to be to arrive at an interpretation. In the end, results from such investigations require verification with more objective paradigms, but they have held up remarkably well (e.g., Haviland and Clark 1974).

In this chapter our focus is on the specific question of how the interpretation of a semantic representation is ordered in relation to the assignment of a syntactic analysis. The corresponding question for introspection is: Do we make judgments about meaning before we make judgments about the structure, or do we access intuitions about these levels of representation in the opposite order? Trying to answer this question by direct introspection is like trying to use introspection to answer any question about internally interdependent processes. In vision, for example, do observers recognize an object in a scene first by isolating its contours and then recognizing what it is, or do they first formulate a hypothesis about what it is and then use that hypothesis to check the contours? One cannot ask observers such questions directly. But it is possible to construct indirect questions that probe the underlying issue.

This example from vision is germane to our present concern. In vision, it was long taken for granted that object recognition must follow identification of object contours, since they delimit the stimulus to be recognized. This apparently logical point is like that of the many theorists who argue that the perception of meaning must follow that of syntax because syntax assigns a particular configuration between words and phrases that delimits the domain of the meaning. Yet, while it may seem logically necessary, in the case of vision it is not true. Recent research has shown that a preliminary local hypothesis about what an object is can guide decisions about where the figure-ground contours lie (Peterson 1994).

In the case of vision it is possible to manipulate the salience of local object versus figure-ground cues and study the resulting shifts in reported location of the figure. In studying the corresponding question in sentence processing, it is less clear what manipulations and measures to use. Sentences are superficially highlighted from their background by a combination of location, voice quality, and continuity. Closure around clause boundaries may offer a corresponding way of studying sentence perception in ways similar to vision.

Sentences offer a feature that visual stimuli do not. Unlike visual stimuli, sentence components are presented in a particular order. This feature of sentences offers a possibility of studying the effects of distortions introduced at specified points. The distortions can be either syntactic or semantic errors, thereby creating a method of studying whether syntax or semantics is accessed first.

A strong assumption underlies the use of introspection to bear on processing order of different kinds of information. This assumption is that the order of appearance to the introspective array corresponds to the order of unconscious assignment during initial processing. This assumption is startlingly redolent of failed nineteenth-century introspectionism. But a past failure is not a principled reason to reject a current attempt, so we must examine the venture on its own terms.

Consider how introspection might work on a behavior with an explicit derivation, such as finding one's way from point A to point B. Suppose further that, while every choice point in the path has landmarks, the directions are a mixture of instructions to make particular directional turns (formal instructions, or "quasi-syntax"), and to go toward particular landmarks (instructions with content, or "quasi-semantics"). Finally, suppose that we contrive an experimental situation so that errors do occur (for example, by making either the turn points or the landmarks somewhat ambiguous). The experimental question is, do people process the landmarks or the turn choices first? The introspective method would be to wait until the end of each complete path. At that point, subjects must introspect on whether they made the right sequence. The relevant question is, in trying to figure out if something went wrong, do subjects *prospectively* retrace their path from the beginning, or *retrospectively* work backward from the end? If it is the former, we might conclude that the data also reveal something related to the original processing; but if the search is retrospective, it would be perverse to conclude that the introspective order corresponds to the processing order. Rather, we might conclude that it corresponds to the posttrip salience of different parts of the just-completed trip.

The same is true of introspection on any process that involves a sequence of operations. Depending on the task structure, the introspection can be prospective (first I did this, then that) or retrospective (I just did that, before that I did this). There are corresponding models of introspection about sentences. Is an introspection prospective or retrospective in terms of processing? According to the prospective model, sensitivity to different sentence features proceeds in the same order as they are actually assigned in comprehension. In the retrospective model, sensitivity is governed by how recent, and therefore how salient, different features are.

6.4.1 Predictions Made by LAST

Consideration of LAST suggests that the relative salience of acceptability violations depends on when the probe event occurs. The expected order corresponds to the hypothesized order of computations during comprehension:

• Basic phrase-level segmentation (i.e., pseudosyntax), including assignment of syntactic category and movement of *wh*-argument gaps into their source location—for example, violations like the following:

(117) *The mouse that I like ate a rarely.

• Meaning violations, as in

(118) *The mouse that I like ate a decision.

• Subcategorization violations, as in

(119) *The mouse I like squeaked a cheese.

• Discontinuous syntactic phenomena, as in

(120) *It's the cheese that I like the mouse who ate.

Each of the above sequences becomes definitively ill-formed on the last word. LAST clearly makes opposite predictions about which kind of ill-formedness will be easiest to detect depending on whether the introspective probe occurs during the sentence processing or after it is complete. If the search is prospective, elementary syntactic and semantic violations should be noted before configurational syntactic violations. On the other hand, if the search is retrospective and occurs after the sentence is complete, the surface sequence and related syntactic structures are more salient immediately following the application of the grammatical synthesis and surface examination during syntax checking.

6.4.2 Prospective Search

The experimental literature does not offer systematic exploration of introspection with acceptability violations. Most incongruity-based tasks presented during a sentence involve comprehension, such as effects on word-reading time, or on overall comprehension success or rate. Nicol, Forster, and Veres (1997) used an online lexical choice task that required comparison of alternative sequences for their relative acceptability at each point. This task is called the *verbal-maze* task (see box 6.11).

Nicol, Forster, and Veres explored the power of local number marking on nouns in controlling agreement with the main verb of a sentence. (We described the reading-time data of Nicol and colleagues on number information in section 6.1.3.) They varied the number on both the head noun of a complex noun phrase and the number on an adjunct noun that immediately precedes the verb. Table 6.19 shows sample materials and the results. The decision times for correctly choosing *is* showed that if the immediately preceding noun is plural, it can slow down reading time. However, an immediately preceding singular noun did not slow down the correct choice of *are*.

There are a number of interpretations for this asymmetry. Nicol, Forster, and Veres argue that plural nouns are marked with an explicit morphological unit, which can then be mistaken as dominating the number of the entire noun phrase. Singular nouns lack an explicit morphological number, and hence cannot dominate the inter-

Box 6.11
The verbal-maze task

In the verbal-maze task, two words are presented at the same time, and the subject has to choose the one that best fits with a continuing sentence. For example, the subject reads one line at a time below, and for each line after the first, must select one of two words that best continues the sentence, in this case obtaining *The author of the speeches is here now*.

The	
by	author
boy	of
the	ran
slowly	speeches
is	dog
here	laughed
funny	now

Response times for selecting the appropriate word at any particular point are assumed to indicate the operation of syntactic and semantic constraints at that point. For example, at the point of deciding between *is* and *dog*, the difficulty of selecting *is* reflects both the requirement that a verb appear next rather than a noun, and the number information on the preceding noun (*speeches*). See Nicol, Forster, and Veres 1997.

Table 6.19
Mean lexical choice decision times (ms) for main verb

Is *Targets*	
SS: The author of the speech is here now.	669
SP: The author of the speeches is here now.	739
Mean for *is*	704
Are *Targets*	
PP: The authors of the speeches are here now.	724
PS: The authors of the speech are here now.	724
Mean for *are*	724

Note: S = singular, P = plural; the first letter refers to number information on the head noun, the second to number information on the adjunct noun.
Source: From Nicol, Forster, and Veres 1997, table 1

Box 6.12
The "stops making sense" task

The "stops making sense" task uses subject-paced word-by-word reading combined with an explicit acceptability task. As each word appears, the subject decides whether it makes sense in the sentence up to that point. If the sentence makes sense at that point, the subject presses one key (e.g., *Y* for yes), which records the response and the reading time on that word, removes the word, and displays the next word in the sentence. If the sentence does not make sense at that point, the subject presses a different key (e.g., *N* for no). This task can be interpreted as tapping online momentary introspections about acceptability. See Boland, Tanenhaus, and Garnsey 1990.

pretation of the number on the noun phrase. This interpretation is consistent with the early phase of pseudosyntax in LAST. It is in the early phase of pseudosyntax that explicit function words and morphemes are quickly recognized and used to create segmentation of major phrases. Overt function morphemes can trigger segmentation. Furthermore, a likely template for quickly determining the number of a noun phrase is to look at the number it carries on its final noun. Thus, as we noted earlier in this chapter, a plural morpheme explicitly marks the end of a noun phrase and also the number it carries.

Another example of prospective search in introspecting about sentences comes from Boland et al. 1995. Boland et al. utilized an online word-by-word "stops making sense" task (see box 6.12). They used a range of sentence constructions. The elegance and clarity of their design and results suffice to stand for a much wider range of research, so we go through their study in some detail. Their ultimate goal was to investigate whether sensitivity to semantic factors could precede the full syntactic analysis that supports the semantic analysis.

The first step was to verify that a semantically implausible object is introspectively detected in the object *wh*-position in object relative clauses. The dependent variable is the cumulative number of "no longer makes sense" responses that subjects generate at each point. We present difference scores, subtracting at each point, the score for the plausible (b) from the corresponding implausible cases (a) (from Boland et al. 1995, fig. 1):

(121) *% Implausible − % Plausible*
a. Implausible: Which prize . . .
b. Plausible: Which client . . .
did the salesman visit _ while in the city?
0 14 40 57

The implausibility of *prize* as an object of *visit* produces a 14 percent difference in "no longer makes sense" judgments on the verb. This result makes several points for us. First, it shows that the *wh*-phrase is immediately moved to the object location, even while reading the verb. Second, it shows that the semantic analysis is arrived at very quickly—immediately in some cases. However, the object position also coincides with the end of the clause, so this result could also be consistent with the view that meaning is not determined until that point.

Boland et al. included a second pair of sentences that controls for this and also explores the effect of subcategorization options (from Boland et al. 1995, fig. 1):

(122) *% Implausible − % Plausible*

a. Implausible: Which movie ...
b. Plausible: Which child ...
did your brother remind _ to watch the show?
0 2 14 37

The fact that the implausibility of *movie* is clearly noticed after *remind* suggests that subjects can immediately access the subcategorization frames of the verb. (Note that the "makes no sense" task is too slow to provide conclusive evidence on the immediate use of subcategorization information; see section 6.1.7.) This allows the expectation that *remind* will be followed by a direct object, as in

(123) Which movie did your brother remind *us* to watch?

Thus, the subject suspends the conclusion that the sentence is implausible until reading *to watch*, at which point there is no further choice.

Boland et al. note that this result could be accommodated by a filtering model in which all category-based alternatives are entertained at the verb and then filtered out by subsequent information (Frazier 1987; Mitchell 1989). They rule out this filtering interpretation by using the "filled-gap" effect (Fodor 1978; Stowe 1986) to show that only one alternative structure is chosen even when more than one exists. The filled-gap effect is the difficulty caused by having a noun phrase immediately after an empty *wh*-object position (e.g., *us* in the sentence below).

(124) Which movie$_1$ did your brother watch *t_1 us make t_1.

The presence of a noun phrase in this position can impede comprehension as measured by reading time and other techniques (e.g., Crain and Fodor 1985; Stowe 1986). The filled-gap effect is taken to reflect the immediate attempt to fill the gap with the available filler, even when it turns out to be inappropriate.

Boland et al. verified the filled-gap effect for their introspective task with simple object relatives versus declarative sentences. In this case, we present the accumulated "no sense" responses for the filled-gap sentence minus the second sentence from Boland et al. 1995, fig. 3):

(125) *% Filled gap − % Declarative*

a. Filled gap: Which star did the assistant watch _ . . .

b. Declarative: I wonder whether the assistant watched . . .

0 1

them photograph last week.

6 16

It is striking that there is even a small effect (a 6% difference) on "makes sense" judgments as a function of the filled gap, since in fact the filled-gap sentence is perfectly sensible. This small effect is further evidence for the automaticity of filling an object argument position with an available unattached *wh*. (We discussed evidence for the early filling of *wh*-gaps in section 6.1.6.)

Boland et al. then showed that the filled-gap effect occurs with verbs like *remind*, which explicitly take different kinds of objects. In this case, we subtract the plausible potential *wh*-filler from the implausible potential *wh*-filler case (from Boland et al. 1995, fig. 2):

(126) *% Implausible − % Plausible*

a. Implausible: Which movie . . .

b. Plausible: Which child . . .

did Mark remind _ them to watch?

1 5 12 14

Boland et al. found that the filled-gap effect occurs more strongly with a plausible than an implausible *wh*-object filler. This would not occur on a filtering model, since the plausibility would have no relative effect until the filtering process itself.

As a final examination of the immediate effects of verb subcategorization, Boland et al. examined the interaction of plausibility and dative verbs. Unlike object control verbs such as *remind*, dative verbs can have animate or inanimate objects. Boland et al. used dative verbs to meet the possibility that the effects with object control verbs were strictly mediated by the animate/inanimate variable as opposed to verb subcategorization. A dative verb can have the double-object frame

⟨NP NP⟩

(127) The maid delivered suite 304 some towels.

that encodes arguments in the order goal and patient. A dative verb can also present just the patient in a transitive frame,

⟨NP⟩

(128) The maid delivered some towels.

or it can have the indirect-object frame

⟨NP PP⟩

(129) The maid delivered some towels to suite 304.

that encodes arguments in the order patient and goal. The data appear below (from Boland et al. 1995, fig. 4):

(130) *% Implausible – % Plausible*

a. Implausible: Which suite ...
b. Plausible: Which towels ...
did the maid deliver _ after she was reprimanded?
0 2 17 25

The fact that there is very little effect on the verb itself suggests that subjects immediately adopt the indirect-object interpretation when the direct-object frame is implausible. Boland et al. (1995:794) summarize their results thus far as follows: "The word position where plausibility effects occur ... depends on the argument structure of the verb; when [it] provides only one possible gap site, [it] is posited and interpreted.... If [there are] alternative gap sites, the filler will be assigned the thematic role with ... the most semantic overlap."

With this as background, Boland et al. tackled the question of whether the semantic plausibility effects of a filler can occur before any possible gap is actually encountered. That is, in terms of LAST, they asked whether the semantic analysis can proceed before syntactic assignment. To explore this, they used nonalternating dative verbs like *distribute*, which do not alternate between an indirect-object/direct-object frame

⟨NP NP⟩

and a direct-object/indirect-object frame

⟨NP PP⟩

but rather have only the latter frame, as shown by the following pair:

(131) Harriet distributed the exams to the students.

(132) *Harriet distributed the students the exams.

Boland et al. constructed examples in which there is a plausible and implausible indirect object. The results showed that subjects recognize the implausibility of the sentence well before the actual *wh*-gap location following the preposition.

(133) *% Implausible – % Plausible*

a. Implausible: Which car salesmen ...
b. Plausible: Which uneasy pupils ...
did Harriet distribute the science exams to?
1 3 17 40

The striking fact here is that the introspection that the sentence is implausible emerges well before the syntactic evidence is present to integrate the implausible object with its verb/preposition. These results suggest that comprehenders can access the meaning directly via the subcategorization frame information, a conclusion consistent with LAST.

So far, so good. But what evidence is there for the immediate application of *wh*-object movement to create NVN structures? Boland et al. (1995) report little if any evidence that NVN governs the recognition of implausible sequences. In particular, there were no clear plausibility effects on the verb for object control verbs (*remind*) or dative verbs (*read*), except for the straightforward object relative clause cases. But data from Boland et al. 1989 with similar paradigms show clear reading-time effects that suggest that the *wh*-filler is at first moved to the direct-object position and that the NVN pattern continues to operate, even if masked by the introspective reports themselves. The data below show differences between response reading time to implausible and plausible cases (from Boland et al. 1989, fig. 1; we subtract the plausible reading times from the implausible reading times, measured in ms):

(134) *Transitive verb*
Implausible reading times – Plausible reading times (ms)
a. Implausible filler: Which stone ...
b. Plausible filler: Which star ...
... did the assistant watch _ all through the night?
10 65 170 120

(135) *Object control verb*
Implausible reading times – Plausible reading times (ms)
a. Implausible filler: Which movie ...
b. Plausible filler: Which girl ...
... did the woman remind _ to watch the show?
5 40 30 100

(136) *Dative verb*
Implausible reading times – Plausible reading times (ms)
a. Implausible filler: Which baby ...
b. Plausible filler: Which poem ...
... did the babysitter read _ in a funny voice?
10 55 70 70

Each case shows some elevation associated with the implausible fillers at the verb, even when the structure would not support it. This is consistent with the postulation in LAST that *wh*-movement and the NVN template applies prior to accessing subcategorization information and meaning. These results suggest that *wh*-movement

and the NVN template are subsequently squelched by the subcategorization information. Whether this is limited to paradigms that require explicit word-by-word judgments remains to be seen.

This concludes our discussion of online prospective introspection. The studies indeed lend support to the proposal that during sentence comprehension, introspective tasks reveal sensitivity to various aspects of pseudosyntax, verb categorical structure, and initially hypothesized meanings.

6.4.3 Context and Introspective Search

Considerable attention has also been given to postsentence judgments. In this case, we expect a different pattern of results. After a sentence, the grammar has generated a complete description and the system has checked it against a memory representation of the input. Thus, the syntax and surface structure should be the most salient sentence features.

A study by Forster and Olbrei (1974) involves a well-formedness judgment paradigm that we think can be manipulated to be either prospective or retrospective, depending on the kind of contrasting ill-formed sentences. They had subjects respond positively to well-formed sentences that varied along two dimensions: semantic reversibility of subject and object, and active versus passive construction. Their goal was to investigate carefully the claim started by Slobin (1966) and echoed by others, that computational effects of syntactic operations such as the passive are neutralized in sentences that are semantically irreversible. Slobin showed that irreversible sentences are matched to corresponding pictures equally fast in the passive and active, while reversible sentences show the expected effect of greater matching time for the passive.

(137) *Nonreversible*
  a. The boy liked the ball.
  b. The ball was liked by the boy.

(138) *Reversible*
  a. They boy liked the girl.
  b. The girl was liked by the boy.

Such results suggested that the syntactic component of comprehension might be circumvented by an apparently more direct perception of meaning based on semantics. Indeed, some took this to show that syntactic computation only occurs as a last resort when semantic analysis is indeterminate. Nonetheless, Gough (1966) showed that the same effects occur when the picture is delayed several seconds after hearing the sentence, thereby suggesting that the effects have to do with accessing the meaning for purposes of picture matching rather than online processing. Gough's result

Box 6.13
The sentence-acceptability task

In the sentence-acceptability task, a whole sentence is displayed on a screen. The subject's task is to press one of two keys to indicate whether the sentence is a meaningful and acceptable sentence. A timer starts when the sentence appears, and it stops when the subject presses a key. Distractor sentences are unacceptable for semantic or syntactic reasons. The nature of these distractors will influence the use of different kinds of information for the acceptability judgment. See Haviland and Clark 1974.

had surprisingly little impact, and many continued to believe that Slobin had shown that semantic analysis proceeds first when it can.

In several studies, Forster and colleagues took another tack. They attempted to show that whether or not a sentence is reversible, plausible, or even implausible, the relative processing load is determined by the syntactic structure. Forster and Ryder (1971) compared semantically irreversible-plausible, irreversible-implausible, and semantically anomalous sentences, each with twenty different syntactic structures. In a comprehension paradigm, the complexity effects of variation in the syntactic structure were strongly correlated across all three semantic conditions. This implicates a constant role for syntax, but it does not demonstrate that semantics is irrelevant to comprehension—there could still be an interaction between the two.

To attack this question more directly, Forster and Olbrei (1974) used a sentence-acceptability task (see box 6.13). They created semantically irreversible sentences in two ways, in half the sentences by varying the animacy of the apparent subject, and in the other half by varying the apparent object. Although they do not comment on this part of their design, it manifestly dealt with the possibility that "reversibility" effects are actually only a function of whether or not the patient is the only varied noun, as in Slobin and many other studies of semantic reversibility. To make it possible to vary the animacy of agents, they used "psych" verbs (e.g., *surprise*), which can have both animate and inanimate "agents." (We use *R* to refer to "reversible" and *N* to refer to "nonreversible.")

Actives

(139) Subject-varying (psych verbs)
  a. R: The agent surprised the model.
  b. N: The idea surprised the model.

(140) Object-varying
  a. R: The boy liked the girl.
  b. N: The boy liked the smell.

Passives

(141) Subject-varying
  a. R: The girl was liked by the boy.
  b. N: The smell was liked by the boy.

(142) Object-varying (psych verbs)
  a. R: The model was surprised by the agent.
  b. N: The model was surprised by the idea.

Forster and Olbrei used two paradigms that differed only with respect to the ill-formed distractor sequences. In one study the distractors were ''almost sentences''; in the other they were ''word salad,'' easily recognized as ill-formed.

(143) *Almost-sentences*
  a. *The reporters printed the senator.
  b. *The workers were repaired by the priest.

(144) *Word salad*
  a. *A poor substitute the amused
  b. *The number light under side

Each test sequence was presented all at once on one line. The subjects were asked to respond positively if the sentence was acceptable, and not to respond at all if it was ill-formed in some way. The results appear in table 6.20.

The quality of the fillers had dramatic effects on the decision-time data. First, the almost-sentence contexts elicited much slower overall response times than the word-salad contexts (188 vs. 257 ms/word). Presumably, this reflects the fact that the word-salad distractors reduced the care with which the sequences must be read to weed out the ungrammatical cases. In addition, there were marked qualitative differences. The almost-sentence contexts elicited both the predicted reversibility effect (63 ms) and the predicted effect of passive (290 ms). The word-salad contexts elicited a similar active/passive difference (166 ms), but showed a weak reversal of the reversibility effect itself (significant in one experiment): nonreversible sentences were actually judged more slowly than reversible sentences (by 35 ms).

Consider first the almost-sentence distractor contexts. Sentences in this paradigm require careful consideration of the details, since analysis in terms of lexical categories, or even coherence of phrases, will not weed out the ill-formed sequences. This explains the slow reading/decision times. It also explains why both reversibility and sentence construction have recognizable effects. In LAST, the checking for grammaticality is carried out during the stages of synthesis and comparison, and confirmation of the initially postulated meaning. At that point, the relative difficulty of the passive construction has accumulated as a function of several stages of processing. The passive violates the NVN template at the initial stage, it involves more computations

Table 6.20
Mean sentence-acceptability decision times for correct *yes* responses

	Nature of distractors	
	Word-salad	Almost-sentence
OVERALL MS/WORD	188	257
Total decision times		
Active		
Rev	1030	1436
Nonrev	1065	1361
Rev effect	−35	+85
Passive		
Rev	1196	1709
Nonrev	1230	1668
Rev effect	−34	+41
Passive-active		
Rev	+166	+273
Nonrev	+165	+307
Overall	+166	+290
MEAN REV EFFECT	−35	+63

Note: Rev = reversible/reversibility, nonrev = nonreversible
Source: Word-salad data from table 2, almost-sentence data from table 3, Forster and Olbrei 1974

during the synthesis, and it takes longer at the syntax checking stage. The much more modest effect of reversibility is an unequivocal factor only after the final meaning is confirmed (see below for discussion of why it is neutralized in the early stages of processing).

Many comprehension models are consistent with the results from the almost-sentence context. Since the decision time is quite slow, even theories that postulate an early stage at which meaning is computed can accommodate the continuing difference between passive and active constructions. But the results from the word-salad contexts more uniquely confirm LAST and other theories that postulate an early stage of computing meaning. In the case of LAST, this difference follows from the fact that subjects can make their decision very quickly, based on a check of the initial meaning. If the sequence has a recognizable meaning, it is well-formed in the context of word-salad distractor trials, and later stages may be ignored in the process of making a decision.

Following our theoretical sketch of the earliest phases of processing actives and passives, the initial application of the NVN template is blocked by the morphology

and replaced by the N + be + Pred template. The sequence then is reinterpreted lexically to arrive at the hypothesized meaning for the passive. The active goes directly from NVN to a hypothesized meaning. This difference in initial computation accounts for the quickly appearing relative difficulty of passives in general (albeit with a somewhat smaller effect than in the almost-sentence context, perhaps because that also includes the syntactic regeneration differences).

The surprising effect in the word-salad context is that reversible sentences are easier than nonreversible sentences, not harder. Most theories of comprehension cannot explain this counterintuitive result. The explanation for the relative ease of reversible sentences in relation to LAST lies in how the reversibility was implemented. Consider actives first. They are subject to the application of the NVN template. This goes through for both reversible and nonreversible object-varying sentences. Since there is no strategy that is affected by the animacy of verb objects, there is no difference predicted for those two active cases. However, the subject-varying actives present a different feature. In subject-varying actives the apparent object of psych verbs can be semantically an agent of some kind. This is reflected in the fact that the object must be sentient. If the apparent agent is animate, the sentences are ambiguous as to whether the primary mental activity is on the part of the apparent subject (145) or apparent object (146):

(145) *Subject agent, active*
  a. The agent surprised the model by going "boo."
  b. *The agent surprised the model by going "boo," but he didn't know it.

(146) *Object experiencer, active*
  The agent surprised the model by being so simple, but he didn't know it.

(147) *Subject experiencer, passive*
  a. *The model was surprised at the idea by going "boo."
  b. The model was surprised at the idea by its being so simple.

It is clear that it is the apparent object of (146), not the apparent subject, that is some kind of experiencer. But the simple interpretation of NVN, in which the initial noun is an agent, is consistent with the sequence, reinforced by its animacy. This is sufficient to determine that the sequence is a grammatical sentence when close attention to grammatical details is not needed for the decision.

Psych-verb sentences with inanimate subjects raise several problems for the initial stages of meaning formation. First, the sentences are not ambiguous, but force an interpretation like (146) above on which the apparent object actually is the only active intentional actor in the situation.

(148) a. *The idea surprised the model by going "boo."
  b. The idea surprised the model by being so simple.

The correct interpretation of the preceding example requires blocking the NVN template assignment of agent status to the initial noun. It also requires reconfiguring the object from being assigned patient to being assigned some kind of agent or experiencer status. All these lexical reinterpretations of "nonreversible" psych actives explain why they are actually more complex to judge than the corresponding "reversible" versions. Since the object-varying sentences do not differ, or differ only slightly, the net effect for all active constructions is in favor of reversible actives.

Now turn to the passive constructions. In all cases, the morphology blocks NVN, and they are then interpreted as N + be + Pred:

(149) a. The girl/smell was [liked by the boy].
b. The model was [surprised by the agent/idea].

In the case of simple transitive verbs, there may be an initially stronger attempt to apply NVN with an initial animate noun, before the morphology triggers the N + be + Pred template. However, after that, the lexical mapping takes the noun from the *by*-phrase and labels it as agent in both cases, hence no further difference as a function of reversibility. So, in this case, LAST predicts a slight superiority for nonreversible sentences.

The opposite is strongly the case for psych-verb passives. At first, the NVN template strongly labels the initial animate noun as an agent. Then the morphology triggers the N + be + Pred template, which unmarks the sentence subject as agent and marks it as experiencer of the predicate phrase. At that point, the same lexical mapping leads to a correct interpretation of the reversible sentences with animate *by*-phrase nouns. That is, as in (145) above, the *by*-phrase animate noun is indeed interpretable as an active intentional agent. But this interpretation is blocked if the *by*-phrase noun is inanimate, since in that case, the only likely interpretation is one like (146) and (147), in which the model is now remarked as the experiencer-agent and the *by*-phrase is unmarked as agent and marked as some kind of patient. Thus, there are several additional computational steps involved in comprehending the nonreversible passives, which accounts for the surprising reversal in the effect of reversibility.

The detailed analysis of the passives suggests that nonreversibility with psych verbs can actually lead to several more miscodings and recodings than does reversibility with simple transitive verbs. Thus, in this set of materials, the net difference is in favor of reversible passives being easier to compute.

Those with connectionist proclivities will be tempted to model these effects as a compound of distinct associative connections—for example, initial nouns are agents, the object of psych verbs are agents, and so on. The results can probably be made to come out the same. This makes sense to us, since our point is that in the word-salad context, all subjects have to do is discover *some* meaning to determine grammati-

cality. This allows the process to rest on the initial stages of meaning formulation, which in LAST is largely probabilistic in nature, and therefore naturally simulated in a connectionist framework.

We conclude that Forster and Olbrei's two paradigms actually tap two different stages of comprehension and introspection. The word-salad contexts tap comprehension and introspection during the apprehension of an initial meaning; the almost-sentence contexts tap these processes after completion of synthesis and checking the output of the grammar against the memory of the linguistic input.

6.4.4 Retrospective Search

Other studies of introspection have arguably concentrated more clearly on postsentence processes, examining the different time course of recognizing semantic versus syntactic violations. Consider the following unacceptable examples from McElree and Griffith 1998:

(150) *Syntactic-configurational*
*It was the essay that the writer scolded the editor
[who admired _].

(151) *Syntactic-verb subcategorization*
*It was the essay that the writer knew
[the editor had gloated _]

(152) *Semantic-distant*
*It was the essay that the writer knew
[the editor had amazed _]

The syntactic-configurational error is related to a grammatical sentence:

(153) The writer scolded the editor who admired the essay.

Syntactic-configurational errors violate a constraint that restricts raising an object from an adjunct relative clause across a *wh*-expression. The meaning of McElree and Griffith's syntactic-configurational errors may be clear, but the syntactic structure is not allowed, perhaps because of a universal constraint on movement out of "islands" (Ross 1974).

The syntactic-verb subcategorization error is related to an ungrammatical sentence:

(154) *The writer knew the editor had gloated the essay.

The error here is in the use of a direct object for a verb that cannot ordinarily take one. Again, a meaning might be imputed—for example, "the editor gloated about the essay."

The semantic-distant error also is related to an ungrammatical sentence:

(155) *The writer knew the editor had amazed the essay.

The semantic-distant error violates a semantic constraint in that the intended syntactic structure and meaning are clear, but it postulates the impossible fact that a person had amazed an essay.

McElree and Griffith studied the question of which kind of violation is explicitly noticed more quickly when people are asked to judge whether sentences are acceptable. In a postsentence judgment task, they recorded the time it takes to note the fact that sequences like these are unacceptable (see box 6.13). They found that response times increased in the order syntactic-configurational (1.14 sec), syntactic-verb subcategorization (1.36 sec), and semantic-distant (1.46 sec).

In a second experiment, McElree and Griffith used a response-signal speed-accuracy tradeoff paradigm, as described in box 6.14. McElree and Griffith found that the intercept (δ) increased in the same order as RTs in the earlier experiment: syntactic-configurational (10 ms), syntactic-verb subcategorization (374 ms), and semantic-distant (435 ms). At the earliest interruption point (14 ms), they found that the d' measures were above chance for the syntactic-configurational sequence, but near chance for the syntactic-verb subcategorization and semantic-distant sequences.

McElree and Griffith take it as obvious that the order of discovering the ungrammaticality of the strings corresponds exactly to the original order of processing. Hence, they conclude that their study supports the conventional view that complete syntactic structures are assigned prior to meaning. We think there are reasons for an alternative interpretation that is consistent with LAST, in which representations are formed in the order pseudosyntax, pseudomeaning, real syntax, real meaning.

First, McElree and Griffith required that subjects did not render their judgments until after they were sure the sentence was over, as cued by the response signal. Making judgments after the sentence is complete is maximally likely to inculcate a retrospective process of introspection. Thus, the subject tries to understand the sentence and gets an initial report from the comprehension system that something is odd. Then to be sure, the subject checks it against the sentence to confirm what is in fact odd about it. According to LAST, at the end of a sentence the grammar has just generated a complete syntax. Thus, at the time of the acceptability judgment, syntactic features are most salient. If the most recently assigned representation is the most salient, this accounts for the immediate ease of noticing the surface violation of the island constraints, next the assignment of the subcategorization structure, and finally the entire meaning of the proposition. At the point of the acceptability judgment, the generative system will have just crashed in the case of the syntactic-configurational violation, since no superficially similar sentence can be will formed. The subcategorization violation takes somewhat longer to be sure about, since super-

Box 6.14
The response-signal speed-accuracy trade-off task

The response-signal speed-accuracy trade-off task is designed to overcome interpretative problems that may arise in response-time tasks. The problem is that since faster responses tend to yield more errors, it is difficult to interpret faster responses as indicating conclusively that the relevant information is available earlier. The response-signal speed-accuracy trade-off task solves this interpretive problem by requiring subjects to make their judgment of acceptability at particular times. This procedure explicitly varied speed and allowed clearer examination of accuracy at different speeds of responding (McElree and Griffith 1995).

In McElree and Griffith 1998, the words of a stimulus sequence were presented visually one after the other for a duration of 250 ms per word. After the end of the stimulus sequence, there was a variable processing interval. A 50-ms, 1000-Hz tone was presented 14, 100, 243, 500, 800, 2000, or 3000 ms after the beginning of the final word of the stimulus sequence. Subjects were instructed to make their judgment of sentence acceptability only on hearing the tone, and within 300 ms of the tone. With a one-hour practice session, subjects were able to make their response within the specified time interval. The subjects were then tested in ten 1-hour sessions that each contained a mixture of sentence types and response lags.

In this paradigm, unbiased measures of accuracy (d') are plotted against response time (including the response lag). The d' measure corrects for guessing by combining the rate at which subjects correctly say that an unacceptable sentence is unacceptable ("hits") with the rate at which subjects incorrectly say that an acceptable sentence is unacceptable ("false alarms"). A speed-accuracy trade-off (SAT) curve is fitted to the d' measures at the various response lags according to

$$d'(t) = \lambda(1 - e^{-\beta(t-\delta)}), \quad t > \delta \text{ else } 0$$

This SAT curve initially shows no recognition of unacceptability (i.e., $d' = 0$), followed by a rapid rise in recognition of unacceptability (β), and ending with a leveling off of recognition of unacceptability at a maximum level of accuracy (λ). The time δ at which the SAT curve begins to rise from $d' = 0$ is the intercept, and is taken to indicate the point at which information becomes available during comprehension.

ficially this sequence corresponds to a possible surface form (as in *It was the essay that the writer knew the editor had admired*). Finally, the semantic-distant sequence is recognized as ill-formed, based on returning to attend to the meaning.

Second, the reaction time to identify the ill-formed sequences was slower than the time to identify well-formed sequences. Faster response times for well-formed sequences would not be expected if the recognition of ill-formedness was proceeding as the sentence is understood. With that strategy, well-formed sentences would require an exhaustive search to be sure they are well formed, while ill-formed sentences could be identified as soon as the ill-formed feature is encountered. Thus, online recognition of ill-formedness would produce faster response times for ill-formed sequences.

Since ill-formed sequences were in fact judged slower than well-formed sequences, introspective judgments must not have occurred "online," but after the sentence ended.

Third, the order of recognition of the different syntactic violations does not clearly conform to the order of processing the corresponding syntactic information. In particular, in many models, these sequences contain an early trigger of a search for a gap. In the syntactic-verb subcategorization sentences, the subcategorization of the final verb is encountered immediately on recognizing the verb. Since the sentence ends at that point, marked by a period presented simultaneously with the verb, the subject can immediately identify ill-formedness (because *gloated* cannot take an object). This judgment is computationally more complex at the corresponding point of the syntactic-configurational sequence because the system has to check whether the potential gap position after *admired* is a legal gap for the unassigned filler (*essay*), and it is not. Prima facie, configurational violations should be recognized more slowly than the corresponding subcategorization violations, but the opposite is the case, as predicted by LAST, for tasks that can be carried out retrospectively.

There is a strategic interpretation of the order of noticing ill-formedness in McElree and Griffith's materials. In this view, the comprehension system reports that a sentence is odd, and then the subject searches back through it to check on that. In the case of configurational violations, the subject needs to look only at the last two words (*who admired*) to note that no sentence could end in that way. Subcategorization violations in the same materials require a span of eight words to be sure that they are ill-formed. (Subjects have to scan back far enough to locate the relative pronoun *that* and then conclude that the subcategorization of the final verb cannot take an object.) Semantic violations require checking back to determine whether the head of the *wh*-phrase (*essay*) is inanimate to make sure there is no interpretation possible. This "strategic" interpretation is consistent with LAST, but more mediated than our first interpretation. That is, to be sure about the ill-formedness of the syntactic violations, reference has to be made implicitly to the syntactic possibilities, which will be enhanced if the syntax is most fresh in mind. Counterfactually, if semantic violations were still easier to find, that would have definitely clinched the claim that they are most salient.

6.5 Conclusion

In this chapter we reviewed a wide range of experimental data. We cited several kinds of evidence that we took to demonstrate an early phase of pseudosyntax. Function words are rapidly accessed, and they prime their syntactically relevant content words. Disruptions in processing that occur because of violations of canonical word order suggest that canonical sentence patterns are applied early during comprehension.

Syntactic category and subcategorization information have immediate effects on reading. *Wh*-traces in argument position immediately prime their antecedents.

Recent studies also suggest that semantic information becomes available early on in sentence comprehension. Violations of semantic constraints influence first-pass reading, but violations of syntactic constraints do not. The argument structure of verbs has early effects on reading.

Several studies suggest that certain syntactic properties have late effects on processing. Higher-level constraints on movement do not influence speeded sentence matching, though they do influence off-line judgments of acceptability. In addition, properties of pseudosyntax and meaning do influence speeded sentence matching. Major phrase and word boundaries attract judgments of click location, but minor phrase boundaries do not, at least immediately. NP-traces prime their antecedents but only relatively late.

We reviewed several studies that suggest that access to semantic and syntactic information depends on when the access task is performed. If it is performed online, semantic information is available more rapidly than syntactic information. If it is performed after a sentence has been understood, syntactic information is accessed more quickly.

Of all the experimental results we have presented in this chapter, there is no single result that critically confirms LAST over all the other theories we reviewed in chapter 4. In fact, there are different theories that can explain various subsets of the results we have cited. For example, the results on rapid identification of function words are compatible with structural theories such as Crocker's or Frazier's, and with constraint-based theories such as Tanenhaus's or MacDonald's. The evidence on *wh*-trace is compatible with theories that adopt an active-filler strategy. Several theories probably could account for the data on introspection by making use of different decision criteria. There also are individual theories that can account for much of the data we have cited. Other than LAST, however, no current theory can directly explain all these facts. There may be a way to amend one of the current theories we reviewed to account for all the results in this chapter. But no theory other than LAST can do so in a way that follows naturally from the architecture of the theory.