7
The Functions of Mental Nodes

> The ultimate function of all neural analysers of sensory input is not mere description or classification, but the shaping of conditional readiness to reckon with the state of affairs betokened by that input. The main question to be answered by the sensory system is not 'What is it?' but 'What does it signify for me?,' or if you like, 'So what?'
>
> (D. G. MacKay, 1984, p. 262)

> The message received may be subject to perceptual errors and confusions resulting from environmental noises, unfamiliarity with the speaker's pronunciation and style, false expectations of the speaker's intent, unintentional ambiguities in what the speaker is saying. . . .
>
> (Warren, 1982, p. 177)

Previous chapters have advanced various sources of evidence for the existence of nodes that play a role in both perception and action. Chapter 7 examines the functional issues. Why have mental nodes evolved? What functions do mental nodes serve?

I examine two possible answers to these questions. One concerns structural economy. Mental nodes are multipurpose processors, and using nodes for more than one purpose may serve to economize on the number of nodes. As we will see, this answer turns out to be surprisingly weak. The other answer concerns the integration of different types of information and turns out to be much stronger. I argue that mental nodes have evolved to enable the rapid integration of heterogeneous sources of information, not just across perception and action but across a variety of other sensory and conceptual modalities. I then examine some of the costs and benefits of this rapid integration process, for example, errors in perception and action (cost) and the automatic resolution of ambiguity (benefit).

The Structural Economy Hypothesis

Two types of economy must be distinguished when addressing the economy issue: structural economy and processing economy (see also Collins & Quillian,

The functions of mental nodes and mirror neurons. Ch 7 (pp. 126-140) in MacKay, D.G. (1987). *The organization of perception and action: A theory for language and other cognitive skills* (1-254). Berlin: Springer-Verlag.

1969; Grossberg, 1982). Structural economy refers to the number of nodes that play a role in some activity, whereas processing economy refers to the number of processing operations that these nodes participate in and how fast these processing operations can be carried out.

Under the structural economy hypothesis, using the same nodes for more than one purpose economizes on structure. Fewer nodes and connections between nodes are required to carry out any given function. In the case of mental nodes, the structural economy hypothesis would be true by definition if perceptual processes were simply the reverse of the corresponding production processes and if no additional mechanisms were required to prevent interactions between top-down and bottom-up processes involving identical nodes.

Unfortunately, both of these prerequisites to easy acceptance of the structural economy hypothesis are false. Chapter 6 presented theoretical and empirical arguments showing that perceptual processes are not simply the reverse of the corresponding production processes, and Chapters 8 through 10 will present convincing evidence that at least one additional mechanism (self-inhibition) is necessary for preventing unwanted interactions between top-down and bottom-up processes involving mental nodes. Without self-inhibition during production, bottom-up priming can cause inadvertent reactivation of higher level mental nodes. Moreover, all but the lowest level perception–action nodes require this self-inhibitory mechanism, a fact that makes the structural economy argument difficult to evaluate and sustain. Mental nodes incur a hidden structural cost and must have evolved for other reasons besides structural economy.

Processing Economy: Integration of Heterogeneous Information

The other type of economy is processing economy. Mental nodes may economize on the rate of processing or on the processing steps required for perception and action. In what follows, I argue that mental nodes automatically integrate heterogeneous sources of information, not just for perception and action but for different sensory modalities such as vision and audition and for different sources of information originating within more cognitive systems (see also Morton, 1969). I argue that, by economizing on processing, and by automatically resolving perceptual ambiguities, broadly defined, this automatic integration process has contributed to the evolution of mental nodes.

The Integration of Perception and Action

That perception becomes integrated with action is self-evident, because perception carries out monitoring functions that are required for the regulation of action (Chapter 9). How and to what extent perception becomes integrated with action is not self-evident, however, and represents a central topic of this book. In the present chapter, I show how mental nodes enable a complete and rapid mesh between perception and action in general, and in Chapter 9, I show how mental nodes facilitate the processing of perceptual feedback during ongoing action.

In the case of mental nodes, perception is synonymous with a disposition to respond. When a mental node becomes activated during perception, all of its associated higher level (e.g., proposition) nodes and lower level (e.g., phonological) nodes become strongly primed or readied for activation under the most-primed-wins principle. And because priming is necessary for action, activating a mental node in perception can be considered the first stage in preparation of a response. Mental nodes activated during perception prime a wide range of possible responses via previously formed connections, both bottom-up and top-down. The top-down priming to lower level nodes enables a repetition response, as occurs during shadowing (Chapter 2), and the bottom-up priming to higher level nodes enables more complex, propositional responses. Following node activation during perception, whatever higher level nodes have accumulated most priming from other internal and external sources can be quickly and automatically activated under the most-primed-wins principle to generate the contextually most appropriate response.

Integration Across Sensory Modalities

Besides integrating perception and action, mental nodes integrate information across sensory modalities. Events in nature usually stimulate more than one sensory system, and mental nodes enable the rapid integration of this correlated information so as to optimize activities within the environment. A typical example is the integration of optical cues to depth with proprioceptive cues originating in the muscles responsible for changing the vergence of the eyes and altering the focal curvature of the lens (Steinbach, 1985). As Warren (1982) points out, these proprioceptive cues "are integrated with, and are indistinguishable from, the purely optical cues to depth. For a person perceiving an object at a particular distance, these (proprioceptive) cues are as fully visual as those providing information via the optic nerve" (p. 189). Interactions between the sight and sound of moving lips during speech perception (the McGurk effect discussed below and in Chapter 2) illustrate another functionally useful integration of information originating in different sensory modalities.

Visual Dominance in Cross-Modality Integration

Mental nodes not only combine information from different senses, they determine the perceptual outcome that results when different sources of sensory information conflict. Mental nodes therefore form an integral part of the explanation of why visual information generally dominates, and determines the nature of the resulting experience, when visual inputs conflict with inputs from any other sensory system. I explore the possible role of mental nodes in three types of visual dominance effect: visual dominance in maintaining balance, localizing inputs, and perceiving speech.

Visual Dominance Over Vestibular Information

Vision dominates in cross-modality integrations with vestibular information (Graybiel, Kerr, & Bartley, 1948; D. N. Lee & Lishman, 1975), and the unanswered question is why. An interesting possibility under the node structure theory focuses on the relative number of low-level nodes devoted to vision versus vestibular sensations. Because retinal neurons greatly excede vestibular neurons in number (Mittelstaedt, 1985), it can be argued that mental nodes receive greater convergent priming from visual inputs than from corresponding vestibular inputs. As a result, when visual and vestibular information conflicts, visual input will contribute more priming, and thereby determine which mental nodes become activated and give rise to perception. Dominance effects between different sorts of information originating within a sensory modality—for example, sacular versus utricular information within the vestibular system—can likewise be explained in terms of relative number of converging neurons (Mittelstaedt, 1985). However, this "spatial convergence" argument can only be applied to low-level, highly automatic perceptual processes. Because the principle of higher level activation is flexible (within limits, Chapter 4; and D. G. MacKay, 1987), the nervous system can readily modify its weighting of different kinds of evidence at *higher levels* (see also Grossberg, 1982).

Visual Dominance in Spatial Localization

Vision also dominates over other sensory modalities in perceiving the spatial location of inputs. For example, when corresponding visual and auditory inputs arrive from conflicting spatial locations, vision usually determines perceived localization. An example is the "ventriloquism effect." Sounds from a concealed loudspeaker can be displaced up to 20 degrees from a visible sound source (e.g., an actor's moving lips), but subjects continue to perceive the sounds as coming directly from the visual source (Witken, Wapner, & Leventhal, 1952). This and many other visual dominance effects are usually explained in terms of the relative reliability of vision versus audition. Under this hypothesis, we tend to rely more on vision than audition, because vision is more accurate than audition for spatial localization. However, some interesting exceptions to visual dominance suggest that this explanation is incomplete or inadequate. For example, neither vision nor audition dominates when the sound track of a film is out of synch or poorly dubbed. We both perceive and are bothered by the asynchrony between visual and auditory events (Neisser, 1976). If people can simply tune audition out during spatial localization, why can't they tune audition out in the case of the poorly synchronized sound track?

The node structure theory provides a more complete explanation of both visual dominance and its exceptions. Phrased within the present framework, the basic issue is this: Why do visual connections to mental nodes generally contribute more priming than auditory or tactile connections? I will discuss two possible answers to this question. One is that for some types of inputs, visual connections have received greater prior practice than connections from other sensory systems (see the discussion of visual skill in D. G. MacKay, 1987). The other is that mental nodes receive greater convergent priming from vision than from any other sensory system. There are simply more nodes in the visual system that can contribute priming, and the convergent priming from these visual nodes is usually simultaneous and continuous in nature, unlike the priming from, say, auditory nodes, which is usually sequential and discontinuous. As a result, visual nodes contribute more spatially and temporally summated priming than do auditory inputs, thereby dominating perception under the most-primed-wins principle.

These hypotheses predict that exceptions to visual dominance will arise when visual and auditory sources have received comparable or asymptotic degrees of prior practice, are comparably discontinuous or sequential in nature, and produce convergent priming from the same number of nodes. Such is surely the case for the bothersome asynchrony between auditory speech and visual lip movements for a poorly synchronized sound track. Both the auditory and the visual inputs arising from the moving lips are discontinuous and sequential, and the perception of timing for visual and auditory speech events probably engages identical timing nodes (Chapter 5). In addition, prior effects of practice for both visual and auditory representations of speech sounds are undoubtedly asymptotic (D. G. MacKay, 1982). The absolute amount of practice for visual and acoustic speech events is so great that the added practice that audition receives through use of, say, telephones and radios provides no additional benefit. As a result, mental nodes will receive nearly equivalent priming from visual and auditory representations of speech sounds, so that neither vision nor audition can dominate under the most-primed-wins principle. The out-of-synch sources will engage in continuous and bothersome conflict with no possibility of automatic resolution.

The fact that timing nodes for vision and audition are shared also predicts that visual dominance will evaporate in a task where subjects estimate the duration of visual and auditory signals, presented either separately or together. Under the node structure theory, dominance depends on the detailed nature of the task being performed and not on fixed dominance relations between different input systems or on general attentional strategies (as in Posner, 1978).

Visual Dominance in Speech Perception

When observing someone speaking, vision dominates when what we hear and what we see conflict in content (rather than in timing). For example, when subjects watch a video recording of a speaker producing a visually distinctive syllable such as *pa* while hearing the syllable *ta* dubbed in synchrony onto the sound track, they usually report hearing *pa* (McGurk & MacDonald, 1976). This visual effect is both unconscious and automatic. Subjects are unaware that the *pa* they "*heard*" originates in the visual signal, and they are unswayed by either instructions or personal experience to the contrary. That is, the illusion persists even when subjects are informed that the auditory and visual inputs differ, and even when they close and open their eyes and observe (to their surprise) that their perception alternates between the visual alternative (with the eyes open) and the auditory alternative (with the eyes closed).

Why does vision dominate in this case, but not in the case of the out-of-synch sound track? As discussed, the fact that we see and hear speech sounds produced with asymptotic levels of practice rules out an explanation in terms of the relative frequency of visual versus auditory "speech events." The fact that both the auditory and the visual inputs arising from the moving lips are discontinuous and sequential also rules out an explanation in terms of temporal convergence of priming. This leaves the spatial convergence hypothesis (discussed previously for vestibular inputs), in which visual nodes outnumber acoustic analysis nodes, so that when the two content sources conflict, visual inputs contribute more convergent priming than auditory inputs and thereby determine which mental nodes become activated and give rise to perception. That is, when a segment is visually distinctive as in McGurk and MacDonald (1976), nodes representing effects of lip movements within sensory analysis systems for vision and audition both converge bottom-up on a set of phonological nodes; but because of their greater numbers, visual nodes contribute more priming, and determine which phonological node receives most priming and becomes activated under the most-primed-wins principle.

Integration Within a Sensory Modality

Mental nodes also integrate different sources of information originating *within* a sensory–cognitive module. For example, the Stroop effect illustrates how mental nodes integrate two sources of visual information: color and orthographic form. Subjects in the Stroop task must name the color of ink that a color word is printed in, and their reaction times are especially fast when color (e.g., "red") and word (*red*) are identical (Keele, 1973). The reason is that both inputs prime the same lexical content node, *red*(color adjective), which is therefore readily activated under the most-primed-wins principle. However, errors and reaction times increase dramatically when the color and the color word differ, because the two sources of visual input conflict and prime different mental nodes in the same (color adjective) domain, the prototypical situation for the occurrence of errors under the theory.

Integration Across Cognitive Modalities

Language is an integrative module par excellence, and can be seen to join cognitive modalities as well as sensory and motor modalities per se. When we say that we like what we see, hear, smell, taste, feel, or any combination of these, language becomes a common source of conceptual integration that spans these sensory–conceptual sources. The way that the smell or sight of apples virtually immediately can evoke the name *apple* also illustrates how language integrates concepts from different sensory–conceptual sources. Experiments on ambiguous figures further illustrate how quickly speech and visual concepts can become integrated. Presenting a word such as *duck* has extremely rapid effects on the visual interpretation of an ambiguous figure such as Jastrow's rabbit-duck (Leeper, 1936).

Mental nodes are the basis for these rapid cross-modal integrations and also enable efficient integration of concepts within the language module. When interpreters translate on-line between languages, for example, they rapidly integrate one set of language concepts with another. A similar integration process occurs at a lower level during shadowing, in which mental nodes rapidly integrate acoustic speech sounds with corresponding muscle movements for producing them. Studies of expert shadowing (e.g., Marslen-Wilson, 1975) demonstrate how rapidly this particular sensory–motor integration can occur. Trained shadowers can repeat back words in sentences with lags of less than 300 ms and begin to reproduce a long word after hearing only its first few segments. Top-down semantic and pragmatic information can also enter into the integrations that occur during "close shadowing." Close shadowers "repair" intentional mispronunciations of the first segment of an input word, reconstructing the original word top-down on the basis of semantic and pragmatic context (Cole & Scott, 1974). These short-lag repairs suggest that semantic and pragmatic processing has an extremely rapid, on-line effect during word recognition, and mental nodes provide a mechanism for explaining how such high-level information can be brought to bear so quickly.

McLeod and Posner (1981) argued that the auditory–vocal integration that occurs in shadowing reflects a "privileged loop" that makes phonology special. In their experiments, shadowing the phonology of a word such as *high* (stimulus) – *high* (response), enabled interference-free dual task performance, whereas producing semantic associates such as *high–low* did not. Under the node structure theory, mental nodes are the mechanisms underlying privileged loops, and it should be possible to observe similarly privileged loops in other skills, levels of language, and aspects of skill. For example, a formally identical integration process occurs at sentential levels during skilled on-line translation between languages. Shared mental nodes provide a privileged loop that enables rapid integration of language 1 input with language 2 output at the sentential level (D. G. MacKay, 1981; 1982). Timing nodes likewise form part of a privileged loop, enabling interference-free execution of many concurrent activities, as when we march, sing, and breathe in time with the band (Chapter 5). Other skills such as expert transcription typing and expert Morse code also involve mental nodes (D. G. MacKay, 1985) that form the basis for other privileged input–output loops under the node structure theory. Indeed, privileged loops (mental nodes) may provide the underlying basis for all highly compatible input–output relations, such as a finger-press response to tactile stimulation of the same finger.

Automatic Resolution of Ambiguity: A Benefit of Integration

To summarize the chapter so far, mental nodes rapidly integrate many different types of information, with perception and action being only the most notable examples, and this integrative process may have contributed to the evolution of mental nodes. However, the automatic integration of information via mental nodes has two inadvertent side effects. I have already mentioned one (visual dominance), and I discuss the other (errors) in the next section. Here I examine how mental nodes disambiguate inputs by rapidly and automatically integrating huge amounts of heterogeneous contextual information. I will argue that contextual resolution of ambiguity is so prevalent and so important as to alone justify the evolution of mental nodes.

Ambiguity can be said to occur when different nodes in the same domain simultaneously receive comparable levels of priming from bottom-up connections. Because studies to date have focused mainly on sentential ambiguities, I begin by discussing examples within the sentential system, but as we will see, ambiguity is a much more general and ubiquitous issue, applying to any node in any system.

Understanding the processing of ambiguity has been especially relevant for attempts over the past several decades to develop machines that can comprehend printed language. These efforts have often been thwarted by the fact that common words frequently allow ten or more distinct meanings (Kuno, 1967). Given this prevalence of ambiguity, what is surprising from the perspective of artificial intelligence is not that people *sometimes* experience difficulty with ambiguity (as shown in D. G. MacKay, 1966), but that they experience difficulty so rarely. The reason lies in the remarkable human capacity to resolve ambiguities by rapidly integrating different types of contextually specified information.

The Contextual Resolution of Ambiguity

Two basic characteristics of disambiguation must be explained in theories of perception: the extremely efficient use of context in disambiguation and the either–or resolution of ambiguity, which is the fact that we perceive either one interpretation of an ambiguous input, or the other, but not both at once (McClelland et al., 1986).

To illustrate the disambiguating effects of context, consider the ambiguous word *crane* and two of its meanings: *crane 1*, a bird with long legs, and *crane 2*, a mechanical hoist. Listeners quickly comprehend one meaning or the other for the word *crane* on the basis of context, including linguistic context, discourse context, situational context (specified via any sensory input or combination of inputs), and beliefs or general knowledge about the topic under discussion. The disambiguating contexts can either precede or follow the ambiguous word and can resolve the ambiguity within 700 ms (Swinney, 1979) or less (D. G. MacKay, 1970d). How can humans bring such large amounts of heterogeneous information to bear so quickly in resolving ambiguity? The contextual resolution of ambiguity constitutes a fundamental problem for artificial intelligence and for theories of speech perception alike.

Ambiguity and the Most-Primed-Wins Principle

A simple, multipurpose mechanism that can use any type of contextual information to automatically resolve any type of ambiguity is built into the node structure theory: the most-primed-wins principle. Figure 7.1 details the process for the ambiguous word *crane*, where a single phonological node, *crane*(stressed syllable), sends bottom-up connections to two lexical content nodes: *crane 1*(noun), which represents the meaning "a bird with long legs," and *crane 2*(noun), which represents the meaning "a mechanical hoist." This is a prototypical instance of ambiguity, because both *crane 1*(noun) and *crane 2*(noun) simultaneously receive comparable levels of priming from *crane*(stressed syllable) when listeners hear the word *crane* in isolation.

However, *crane 1*(noun) and *crane 2*(noun) are unlikely to achieve exactly equal priming in everyday speech perception, because a large number of additional (contextual) sources will be contributing additional priming to one node or the other. For example, *crane 1*(noun) will receive priming from contextual sources representing, say, visual perception of the bird, a discourse on cranes or on birds in general or a prior sentential context such as "The birds with the longest legs, biggest beaks, and greatest wingspan in swampy habitats are. . . ." In such contexts, *crane 1*(noun) will achieve greatest priming and become activated automatically under the most-primed-wins principle, rather than *crane 2*(noun), referring to the mechanical hoisting device (D. G. MacKay, 1970d).

Because disambiguating contextual information is almost invariably available during everyday conversations, the theory explains why ambiguity causes so little trouble in real life. Context normally ensures that the appropriate node receives more priming than any other node in its domain, which in turn ensures that the appropriate meaning becomes perceived. Ambiguity can only cause

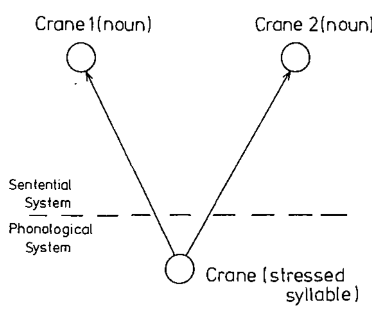

FIGURE 7.1. An illustration of ambiguity in the node structure theory. The syllable *crane* and two of its connected lexical content nodes are shown.

problems in the theory if the intended meaning is contextually inappropriate (as in garden path sentences, Chapters 2 and 4), or if both nodes representing the ambiguous meanings are receiving exactly equal priming, not just from below, but from all currently available contextual sources (D. G. MacKay, 1970d). Needless to say, these two conditions seldom arise in everyday life.

Conceptual Frequency and Context-Independent Disambiguation

Although normally sufficient, context is unnecessary for resolving ambiguities under the node structure theory, because the most-primed-wins principle can disambiguate words on the basis of conceptual frequency, even when the contextual cues that normally predispose perception of one meaning rather than the other are absent. Experiments in which an ambiguous word such as *crane* is presented in isolation nicely illustrate this point (Hogaboam & Perfetti, 1975). Here frequency of prior activation of the nodes representing the two meanings will determine which meaning becomes perceived, because frequency influences linkage strength, degree of priming, and probability of activation. Subjects will tend to perceive whatever meaning has higher frequency of occurrence in their personal experience. All other factors being equal, the ornothologist will perceive *crane 1*, whereas the hoist operator will perceive *crane 2*.

Averaged across subjects, of course, frequency of personal experience correlates with frequency in the language, and this explains a wide range of phenomena in the literature. To pick just one relevant example, subjects presented with a "ditropically ambiguous" expression, such as "He kicked the bucket," first perceive the high-frequency, idiomatic meaning, "to die" (Van Lanker & Carter, 1981). The literal (nonidiomatic) meaning, "a bucket was kicked," is less frequent than the idiomatic meaning and therefore less likely to become activated under the most-primed-wins principle.

The Either–Or Resolution of Ambiguity

The fact that ambiguities are resolved on an either–or basis imposes another basic constraint on theories of perception. Why do we initially perceive *only one* interpretation of the Jastrow rabbit-duck, or comprehend *only one* meaning of an ambiguous word or sentence (Kahneman, 1973; D. G. MacKay, 1966; 1970d)? In the example under consideration, why do we perceive either *crane 1* or *crane 2*, but never both simultaneously? Similarly, why do we comprehend an ambiguous sentence such as "They are flying planes" to mean roughly either, "Those machines are planes that fly!" or "Those people are in the process of flying planes!" but (almost) never *both* meanings simultaneously?

Unlike other theories, such as that of D. G. MacKay (1970d) and McClelland et al. (1986), the node structure theory requires no special-purpose mechanisms such as reciprocal inhibition between content nodes for accomplishing either–or resolution of ambiguity. The most-primed-wins mechanism, which is required for other reasons in activating each and every node, automatically resolves ambiguity in an either–or way. Under the most-primed-wins principle, only the most primed node becomes activated in a domain, including the domain of sentential sequence nodes. Thus, with presentation of an ambiguous sentence such as "They are flying planes," either COPULA or COMPLEX VERB becomes activated, and as a consequence, either *are*(copula) or *are flying*(complex verb) becomes activated, but not both at once (see McClelland & Kawamoto, 1986, for a similar account involving "case-role" units instead of sequence nodes or syntactic category units).

Of course, if instructed to do so, subjects *can* perceive first one and then the other meaning of an ambiguous word or phrase. However, perceiving the second meaning takes considerable time (D. G. MacKay & Bever, 1967), because a nonautomatic process is required to boost the priming of nodes representing the second meaning so that these primed but not activated nodes can become activated under the most-primed-wins principle when the activating mechanism is applied again. Interestingly too, the time required to perceive the second meaning is even longer when a *different* activating mechanism must be applied, as when the two meanings of the ambiguity belong to different domains or syntactic categories such as *like*(verb) versus *like*(preposition) (D. G. MacKay & Bever, 1967).

The Time Course of Disambiguation

Swinney's (1979) experiments provide a clear picture of the time course of disambiguation in sentence comprehension. Subjects listened to sentences such as "Rumor has it that for years, the government had been plagued with problems. The man was not surprised when he found several spiders, roaches, and other bugs in the corner of the room." Immediately after hearing the ambiguous word *bugs* in this passage, subjects saw either a word or a nonword on a screen they were watching, and made a yes–no lexical decision as quickly as possible. Swinney found that up to 400 ms following *bugs*, lexical decisions for words related to either of its meanings (e.g., *spy* and *ants*) were faster than for unrelated control words. The reason is that the node *bugs*(syllable) primes both *spybugs*(noun) and *insectbugs*(noun), thereby facilitating lexical decision time for *both* sets of semantically related words. However, *insectbugs*(noun) also receives contextual priming from *spiders*, *roaches*, and so on, and because *insectbugs*(noun) has more priming than *spybugs*(noun), it becomes activated under the most-primed-wins principle and enters conscious awareness about 700 ms after the syllable *bugs* is heard. Meanwhile, the original priming of *spybugs*(noun) has decayed, so that only lexical decisions for words related to the consciously perceived meaning (*ants*) receive facilitation 700 ms after *bugs*.

Ambiguity, Nonunique Priming, and Shades of Meaning

Ambiguity can be considered a special case of the more general phenomenon of nonunique priming, discussed in the previous chapter. Like ambiguity (Figure

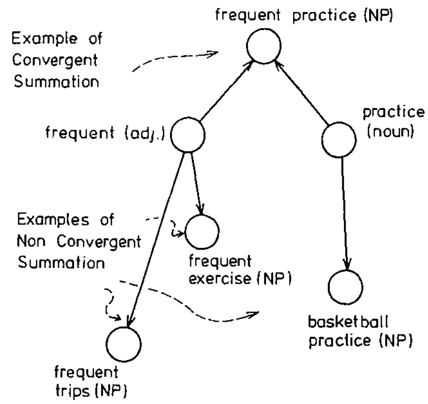

FIGURE 7.2. An illustration of nonunique priming, where four nodes in the same domain (Noun Phrase or NP) receive simultaneous first order priming. Note that priming converges or summates spatially for only one of these nodes receiving nonunique priming.

7.1), nonunique priming occurs whenever two or more nodes in the same domain become primed at the same time, usually from below. The difference is that in the case of ambiguity, nonunique priming arrives at comparable levels and from exactly the same number of bottom-up connections. To illustrate this difference, Figure 7.2 takes a typical (unambiguous) example of nonunique priming, the lexical input *frequent practice.* Activating the lexical nodes for *frequent* and *practice* primes a large number of noun phrase nodes at the same time. In particular, activation of *frequent*(noun) nonuniquely primes noun phrase nodes such as, say, *frequent trips*(noun phrase), and *frequent exercise*(noun phrase), as well as *frequent practice*(noun phrase) (Figure 7.2). Likewise, activation of *practice*(noun) nonuniquely primes noun phrase nodes such as, say, *basketball practice*(noun phrase), as well as *frequent practice*(noun phrase). However, because of convergent summation, *frequent practice*(noun phrase) will receive more priming than any of these other nodes in its (noun phrase) domain (Figure 7.2) and become activated automatically under the most-primed-wins principle, thereby determining perception. Clearly, ambiguity presents more of a problem for accurate comprehension than do other types of nonunique priming, but the most-primed-wins principle solves both problems in the same way under the theory, that is, automatically, without recourse to a conscious decision process and in a categorical or either–or way.

Interestingly, "shades of meaning" reflect the weaker case of nonunique priming under the node structure theory and therefore differ from ambiguity (unlike the proposal of McClelland & Kawamoto, 1986, p. 315, where ambiguity shades off seamlessly into shades of meaning). To illustrate shades of meaning, the *student* in *student of life* differs from the *student* in *medical student.* However, under the node structure theory, the same node represents *student* in these two examples. *Student*(noun) nonuniquely and nonconvergently primes both noun phrase nodes, *student of life*(noun phrase) and *medical student*(noun phrase). The differing shades of conceptual students are represented by the differing proposition nodes that *medical student* and *student of life* connect to.

Generality of the Problem

Recall that ambiguity can be said to occur in the node structure theory whenever different nodes in the same domain simultaneously receive comparable levels of priming from bottom-up connections. So defined, ambiguity can arise at every level in every system and may occur at least as frequently and can represent at least as much of a problem for phoneme recognition as for word comprehension (see also Massaro, 1981; McClelland & Elman, 1986). If segment nodes representing *g* and *k* receive comparable levels of bottom-up priming, the input can be said to be phonologically ambiguous between *g* versus *k*, for example.

Both experimental and theoretical considerations indicate that ambiguity is relatively common at the phonological and phonetic levels. Experiments such as D. G. MacKay (1978) show that, presented in isolation, words are highly ambiguous phonological entities and are subject to frequent misperceptions (Chapter 2). Moreover, if an easily perceived word in a naturally produced sentence is spliced out and presented out of context, subjects have difficulty telling what the word is (Cutler, 1985). Sentential context is apparently as necessary for deciphering speech sounds as for resolving lexical ambiguities, and the prevalence of phonetic and phonological ambiguities undoubtedly represents one of the main reasons why we have so far been unable to develop computer programs for providing accurate phonemic analysis of spoken English (see McClelland & Elman, 1986).

Theoretical considerations also suggest a prevalence of phonological ambiguities. Sensory analysis and phonological feature nodes generally prime a sizable set of segment nodes at the same time. Assume for the sake of illustration that a single node represents the phonological feature +voice (in either initial or final syllabic position). Activating this feature node would simultaneously prime the entire set of nodes representing voiced speech sounds, over 30 segment nodes in all. This, plus the occurrence of coarticulational overlap between adjacent segments further multiplies the theoretical likelihood of phonological ambiguity (McClelland & Elman, 1986).

Fortunately, however, direct resolution of phonological ambiguities is neither desirable nor necessary during everyday human sentence perception. It is not desirable because lower level disambiguation requires activation of phonological nodes, which would reduce rate of processing. If phonological nodes routinely became activated, the probability of activating the wrong node would also increase, causing perceptual errors (Chapter 4). And resolution is not necessary because the existence or nonexistence of nodes at higher levels usually resolves phonological ambiguities automatically. For example, an acoustic input halfway between *gastrointestinal* versus *kastrointestinal* is ambiguous at the phonological level between the syllables *gas* versus *kas*, between the segments *k* versus *g*, and between the features +voice versus −voice, but is unambiguous at the lexical level. Nonexistence of the node *kastrointestinal*(adjective) eliminates the ambiguity.

The principle of higher level activation reduces the probability of perceptual error in another way as well. Higher level nodes receive disambiguating information that is unavailable to lower level nodes but not vice versa. As noted previously in this chapter and in Chapter 4, lexical content nodes receive first-order priming from external (nonspeech) sources that cannot reach phonological nodes. A lexical content node, such as *apple*(noun), receives convergent priming from nonspeech sources representing, say, the smell or sight of an apple, as well as from phonological nodes representing the word *apple*. These external sources of priming can therefore serve to disambiguate an input at the lexical level but not at the phonological level. Phonological nodes representing the segments *l*, *p*, and *a* of the words *apples*, *pals*, or *laps* do not receive direct connections from the visual nodes representing *apples*, *pals*, or *laps*. And because phonological nodes are not subject to external sources of conceptual disambiguation, routine activation of phonological nodes (contrary to the principle of higher level activation) would further increase the probability of misperception.

Errors: A Cost of Integration

Mental nodes and the way they integrate information incur both costs and benefits. If integration of heterogeneous sources of information and automatic resolution of ambiguities represent benefits of mental nodes, errors represent a cost. In what follows, I review the general characteristics of errors discussed in previous chapters in order to show how automatic integration of heterogeneous sources of information via mental nodes contributes to errors in perception and action.

The Stroop effect illustrates how integration of different types of visual information can lead to errors (see also Norman, 1981), and it is an interesting (but often overlooked) historical fact that a whole range of Stroop-like effects were originally observed as a type of speech error. Meringer and Mayer (1895) reported that the names of colors and objects that a speaker is looking at, has heard spoken, or has recently read, often intrude as speech errors, substituting for a word that the speaker currently intends to produce. As is characteristic of speech errors in general, the intruding color words generally belonged to the same syntactic category (adjective), and subcategory (color adjective), as the word that the speaker intended to say at the time. Such errors illustrate how lexical nodes integrate priming that arises from a visual color, with priming that arises from reading, hearing, comprehending, and producing other (syntactically similar) words. Stroop errors occur because the wrong source of priming happens to dominate at the time when the activating mechanism is applied, so that the wrong lexical content node becomes activated under the most-primed-wins principle.

Blends and phonologically similar word substitutions also illustrate how errors arise from the integration of top-down and bottom-up priming during speech production. When speakers substitute words that are syntactically and phonologically similar, such as *a pressure* for *a present*, the syntactic similarity (both are nouns) can be characterized as a top-down effect, while the phonological similarity reflects a bottom-up effect (Chapters 2 and 6). "Freudian" errors also illustrate how mental nodes integrate top-down and bottom-up information of a much more heterogeneous sort during speech production (Dell, 1980; D. G. MacKay, 1982). An example is the substitution of *battle scared* for *battle scarred* in reference to an army officer whom the speaker believes is *scared of battle*. Although the speaker wishes to keep this opinion secret, top-down priming from this currently active "propositional belief" nevertheless automatically influenced which node in the (past participle) domain becomes activated. Another currently active belief also influenced the speaker's "correction" of the error, "battle scared, excuse me, I mean bottle scarred. . . ." Under the Freudian analysis (Freud, 1901/1914), this new error reflects an additional belief that this *battle scared* officer has been "*hitting the bottle*."

Freudian slips of the ear illustrate similar integrations of top-down priming (arising from propositional beliefs and attitudes) with bottom-up priming during ongoing word perception. An example is the misperception of *carcinoma* for *Barcelona* in the case of an individual who is temporarily concerned or preoccupied with this particular disease. The misperception occurs because priming for *carcinoma*(noun), arising from the preoccupation (top-down), and from aspects of the acoustic stimulus (bottom-up) exceeds priming for *Barcelona*(noun) arising from the input itself. As a consequence, the extraneous node *carcinoma*(noun) becomes activated under the most-primed-wins principle rather than the intended node *Barcelona*(noun).

In conclusion, errors in perception and action are largely attributable to the automatic manner in which mental nodes integrate priming from heterogeneous sources. However, I will argue in Chapter 9 that mental nodes also make errors especially easy to detect and correct, and this fact, together with the relative infrequency of errors in perception and action, suggests that errors constitute a small price to pay for the benefits of mental nodes, such as the automatic resolution of ambiguity.