CHAPTER 12:

LANGUAGE PRODUCTION

CHAPTER 12:

LANGUAGE PRODUCTION

Preview questions

1. How does language differ from other forms of communication?
2. What do slips of the tongue tell us about the processes involved in speech production?
3. How might brain damage affect speech production?
4. How might we define 'language'?
5. What kinds of processes are involved in writing?

© mrPliskin/iStock

INTRODUCTION

In 1970 in California, a case of child neglect was discovered that was to have a profound impact on our understanding of language development. A girl, named in the literature as 'Genie', had been isolated from the age of 20 months until she came to the attention of social services at the age of 13 years and 7 months. When found, Genie was undersized and severely malnourished. She had painful calluses from being physically restrained over long periods. Throughout her years of isolation, Genie had not been spoken to and she had been beaten when she made a noise. She had not been exposed to language – she had spent many years locked in a room at the back of the family home where she was not able to overhear her family's conversations, and where she was not exposed to language sounds from radio, television or other sources (Curtiss, 1977).

When Genie was found, she did not speak and seemed to understand no more than a few words (Rymer, 1992). Once Genie was taken into care and was exposed to language, in some ways her language development seemed to proceed as it would for a younger child who had typical exposure to language (Curtiss, 1977). She progressed from single words, to two-word and then three-word combinations, and she rapidly acquired vocabulary. However, her language development showed some significant deviations from the normal pattern. Genie had a vocabulary of more than 200 words before she began to combine them, whereas children typically combine words earlier. The word types evident in her early vocabulary also differed from the normal pattern. For example, while most children's early vocabulary consists of basic class words (e.g. 'dog', 'cat'), Genie's vocabulary development showed an emphasis on colours and numbers, shape and size terms, and basic (e.g. 'dog'), superordinate (e.g. 'animal') and subordinate (e.g. 'Labrador') category words (Curtiss, 1981; Fromkin et al., 1974). Genie seemed to seek out words that would allow her to differentiate between similar objects (e.g. pen versus pencil) rather than acquiring labels for a category of object. By contrast, a typically developing child might initially use 'pen' for pens, pencils, crayons and other objects of similar shape, a pattern referred to as over-extension.

Syntax refers to the rules governing the ways words can be combined to create meaningful sentences.

Content words are words that provide meaning to the sentence; these contrast with function words which do the grammatical work of the sentence.

Language production refers to a number of processes by which we convert a thought into language output, in the form of speech, sign language or writing.

Social cognition refers to the ways in which people make sense of themselves and of others in order to function effectively in a social world.

The most striking feature of Genie's language reflects her problems developing **syntax** or grammar (see Table 12.1), evident in the ways she combined words. Curtiss (1981) described Genie's sentences as '. . . the stringing together of **content words**, often with rich and clear meaning but with little grammatical structure' (p. 21). Some examples of Genie's utterances illustrate this feature of her language: 'I like hear music ice cream truck'; 'Think about Mama love Genie'; 'Dentist say drink water'; 'Applesauce buy store' (Curtiss, 1977). Genie's case reflects three key issues in language acquisition. First, her failure to fully acquire language suggests that there may be a critical or sensitive period for language acquisition, and particularly for grammar development; if the child is not exposed to language in a social context within this period, normal development is constrained (see Lenneberg, 1967). Second, her language reflects the dissociation between the acquisition of vocabulary and the flexible use of this vocabulary to form novel sentences. Third, her case suggests that language acquisition, like many cognitive functions, relies on interplay between input from the environment (nurture) and biological makeup (nature).

This chapter examines the nature of language and the cognitive processes involved in **language production**, with a focus on speech. Language is a quintessentially human ability and the ability to communicate our thoughts to others through language is fundamental for **social cognition**. Language also shapes mental representation and thinking (e.g. Crystal, 1997). Once acquired, language is 'fundamental to all distinctly human thought and consciousness' (Donald, 1999, p. 139). It is therefore important for cognitive psychology to study the processes involved in speech production and to try to understand how a thought is turned into spoken words.

Two basic stages in speech production have long been recognized: the formulation of a thought and its conversion into speech (e.g. James, 1890). However, cognitive psychology

Table 12.1 Levels of linguistic analysis

Level	Refers to:
Semantics	The level of meaning in language
Syntax	The rules by which words are combined to make meaningful sentences
Morphology	The rules by which words are constructed and modified
Phonology	The sound units within a language

remains relatively poorly informed as to the precise nature of the processes underlying language production. As we will see in the next chapter, greater progress has been made towards understanding language comprehension, while less research has addressed the processes underlying production (MacNeilage, 1999). Methodological constraints provide one reason for this bias: it is difficult to control experimental stimuli in order to study language production. When examining language comprehension, we can manipulate the words, sentences or other stimuli that are presented to research participants and measure the effect on comprehension. Comprehension follows on the presentation of the stimuli. But production proceeds from cognition to motor output, and it is a far more difficult task to control or inspect the content of someone's thoughts. Speech production proceeds in a top-down manner, that is, it is **conceptually driven** (see Chapter 2).

Conceptually driven or top-down processes reflect the influence of higher order cognitive processes such as thoughts, beliefs and expectations.

In spite of the methodological challenges posed by the topic, our understanding of speech production has seen substantial development in recent years. Knowledge of speech production derives from a number of sources, involving, for example, experimental methods, computational modelling, neuroscientific methods and neuropsychological case studies. In the present chapter we focus on learning about language production by examining what happens when the system fails. Two types of system failure are considered. First, we consider speech errors; slips of the tongue and other speech errors reveal much about the processes underlying speech. Errors can be induced experimentally or recorded from spontaneous speech. Second, the effects of physical damage to the areas of the brain responsible for language production will be examined.

Let's start by considering what language is: How might we define 'language'?

LANGUAGE AND COMMUNICATION

Language is our principal means of communication and forms the basis of the majority of social interactions. **Communication** can be fairly readily defined as any means by which information is shared (e.g. Field, 2003) or as a process whereby 'a source encodes and transmits a signal, which is detected by a receiver and decoded into meaningful terms' (MacFarland, 1999, p. 387). Many definitions of language would include its use in communication as a core feature, but clearly language goes beyond communication. The information sharing function of language may be a relatively minor role; Aitchison (1996, p. 25) suggests that language has been particularly important for human evolution because it promotes social bonds and social interaction and because it provides an effective means of persuading others.

Communication refers to any means by which information is shared.

There are two ways in which we can use language to communicate. One way is through writing. Written language is a new (and arguably humankind's greatest) invention; the earliest evidence of writing dates to about 5000 years ago. Writing developed from a number of distinct systems originating in different parts of the world and this is reflected in the considerable cross-language variation in scripts today (see Chapter 13). Writing involves converting thoughts or speech to print. In contemporary society, reading and writing are essential skills and anyone who fails to acquire them, for whatever reason, is at a considerable disadvantage. Writing also plays a vital role in language survival, by allowing a record of the language to be retained across generations: a language without writing is unlikely to survive.

The principal way we use language is through speech (or a related mode of output such as sign language). This aspect of language has been a feature of human cognition for tens of thousands of years, and is without parallel in the animal kingdom. Spoken language is found in all human groups and would seem to be qualitatively different from the communication systems of other species.

People also communicate non-verbally. Non-language vocalizations (e.g. grunts, groans) can convey information, and gestures can supplement or substitute for spoken language (see Jacobs & Garnham, 2007; McNeil, 1992). There are many speculative accounts of the origins of human speech. Some accounts highlight the interaction between spoken language and gesture use (e.g. Corballis, 2003). Gesture is so closely tied to human language that we continue to gesture even when we cannot be seen; it is common for people to gesture as they communicate over the telephone, for example (Bavelas et al., 2008; see Chapter 8 for more on gesture and embodied cognition). Subtler non-verbal signals such as body language and tone of voice (see Chapter 2) also communicate to others, whether we are aware of this or not.

Languages vary on a number of dimensions, but also have features in common, an issue we explore next.

LANGUAGE UNIVERSALS

Estimates regarding the number of languages in use worldwide vary considerably, depending on the criteria used to count speech systems as distinct languages. There are about 6000 languages (e.g. Comrie 1989; Crystal, 1998; Krauss 1992; Moseley, 2007) in use worldwide, and this figure is decreasing. Many of these languages are close to extinction, with a very small number of speakers and few or no child speakers. For example, Krauss (1992) noted 187 indigenous languages in North America, only 20 per cent of which were still being learned by children. In fact, just 4 per cent of the world's languages is spoken by 96 per cent of the world's population, placing many languages on the 'endangered' list (Crystal, 2000). Krauss (1992, p. 7) estimated that as few as 10 per cent of the world's languages will remain in a hundred years, with minority languages facing increasing pressure from the dominant, majority languages (see also Crystal, 2000).

Languages vary in the number and type of sounds used, in basic word order, in the size of their vocabularies (reflecting the number of items in the **lexicon**) and in their rules for sentence construction. However, all are capable of expressing complex and new ideas: there are no 'primitive' languages. Though the precise way in which concepts are expressed may differ across languages (e.g. Boroditsky, 2001; see Chapter 13), the expression of complex ideas is evident in all languages and in all human groups.

Mental lexicon is our store of knowledge about words and their uses.
Linguistic universals are linguistic features said to be found in all languages.

Languages have some key features in common, though it proves to be a difficult task to identify a set of **linguistic universals**, that is, features that are shared by all languages. Aitchison (1996, p. 177) lists the following features as 'absolute universals' and acknowledges that even these are problematic. According to Aitchison, all languages:

- have consonants and vowels;
- combine basic sounds into larger units;
- have nouns (words for people, places and objects, e.g. book);
- have verbs (words that represent actions or 'doing', e.g. to read);
- can combine words in meaningful ways;
- can express who did what to whom;
- can express sentences as negatives;
- can express sentences as questions;
- are structure-dependent, that is, involve a syntactic structure or grammar;
- allow recursion (the use of a rule within itself, allowing, for example, embedded sentences).

There are, as Aitchison points out, immediate problems with this listing. Sign languages are language, though they do not use a system of vowels and consonants. Some languages do not reliably distinguish between classes of nouns and verbs. Some nouns can represent actions (e.g. 'destruction') and nouns can be used as verbs ('to text'). Table 12.2 provides a summary of the parts of speech, such as nouns and verbs. While all spoken languages are based on combinations of vowel and consonants sounds, the precise set of sounds varies considerably across languages (see the section on phonology to follow). MacNeilage (1999) identifies the syllable, that is a vowel and consonant combination, as a universal unit of speech. Some languages use other sounds in addition to vowels and consonants. For example, in **tonal languages**,

Tonal languages use changes in tone to alter the meaning of the word.

Table 12.2 The lexical categories (word classes or 'parts of speech') of English. See *The Cambridge Grammar of the English Language* (Huddleston and Pullum, 2002).

Category	Description
Verb	A very large class of words among which are many denoting actions (*run*, *jump*, *swim*) or states or experiences (*feel*, *see*, *think*). Auxiliary verbs mark tense, aspect and modality (*has eaten*, *is eating*, *may eat*). Verbs that take a direct object (as in *The dog ate the bone*) are called transitive verbs. Others are intransitive (*The dog slept*).
Noun	An extremely large class of words including the words denoting people, places, things, or abstract ideas in the broadest sense. The nouns include not only *desk*, *book*, *football*, *house*, etc., but also *absence*, *thought*, *possibility*, *failure*, etc. Proper nouns begin with a capital letter and name particular people, places, or things: *Barack Obama*, *London*, *Disneyland*. The pronouns are a special subclass of the nouns, used for referring to things without naming them (often because they are named elsewhere in the context): *he*, *she*, *I*, *you*, *they*, *it*.
Determinative	The fairly small class of determinatives includes words usually used with nouns to specify definiteness or quantity: *the*, *a*, *some*, *all*, *every*, *many*, *much*, *several*, *this*, etc. Sometimes they occur without nouns (*Some may disagree*; *I like this*), or in other uses (*So much the better*).
Adjective	The adjectives are a large class of words that typically modify or qualify the meaning of nouns, as *brown* does in *the brown dog* or *Our dog is brown*: words like *happy*, *cloudy*, *intelligent*, *mysterious*, *cool*, and many others.
Adverb	The great majority of the adverbs are formed from adjectives by adding *ly*: *nice* is an adjective and *nicely* is an adverb. They very commonly modify words other than nouns, providing information about where, when, how, or to what extent things happen: *locally*, *recently*, *awkwardly*, *extremely*, Adverbs unrelated to adjectives include *soon*, *quite*, *too*, *always*, *seldom*.
Preposition	Prepositions often precede noun phrases to indicate spatial or temporal relationships, as in *the book is on the shelf*. The prepositions include *after*, *at*, *before*, *by*, *despite*, *during*, *except*, *for*, *in*, *into*, *of*, *on*, *over*, *since*, *through*, *to*, *under*, *with*, *without*. some occur before clauses (*after we met*), and some occur without anything following (*Go right in*).
Subordinator	There is a very small class of words called subordinators that function to mark clauses embedded in other clauses: *that* as in *that nobody loves me*; *whether* as in *whether anybody loves me*; *for* as in *for someone to love me*.
Coordinator	There is a very small class of words called coordinators, used to link clauses or phrases together: *and*, *or*, *but*, *nor*, and a very few others. (The subordinators and coordinators are traditionally classed together and called 'conjunctions'.)
Interjection	An interjection is a word that interrupts a sentence with an immediate expression of emotion or sentiment: *yikes*, *oops*, *shush*, *ouch*, etc.

altering the tone of expression communicates meaning. In English, and non-tonal languages generally, we can change the tone of the utterance for emphasis, or to convey emotion, but doing so does not alter the meaning of the word. In a tonal language, the tone carries meaning. In Mandarin, for example, a language with a relatively small number of syllables, 'ma' can mean 'mother', 'horse' or 'scold', if the tone is even, falls then rises, or falls, respectively (Ladefoged, 1993, p. 255).

Some languages use unusual classes of consonant sounds. There are about 30 languages in Southern Africa (e.g. the Xhosa, Khoikhoi and Sosotho languages) which use a 'click' sound (like the 'tut tut'/ 'tsk' sound in English). These clicks may originally have aided communication while hunting, as they mimic natural environmental sounds and therefore would not have startled prey (Knight et al., 2003). Sounds that qualify as meaningful differ substantially across languages. Even in languages with many features in common (e.g. German and English) the precise set of speech sounds varies. Given this diversity, an approach to defining language based on broad design features may prove more fruitful.

Hockett's design features for human language

Scan to watch a video about animals learning language

Charles Hockett's (1960) set of 16 design features for human language was formulated with the aim of identifying properties that are unique to human languages and differentiate them from other animal communication systems. While animal communication systems share some of the following features, only human language demonstrates the full set. The design features treat speech as the standard mode of expression.

1 *Vocal–auditory communication channel*: Languages normally transmit information via spoken sound with the sender speaking and the receiver hearing the spoken signal (sign systems use the tactile-visual medium to similar effect). Many animal systems use vocalizations as the means of communicating while other animals use non-vocal means to communicate. For example, a honeybee uses a 'figure of eight' shaped dance to signal information about the location, distance and quality of a food source (Von Frisch, 1962).

2 *Broadcast transmission and directional reception*: The speech signal is transmitted out from the source (the speaker's mouth) and is localized in space by the receiver.

3 *Rapid fading*: The spoken message fades after production, unlike, for example, written language, which can be inspected over time.

4 *Interchangeability*: The sender can also be a receiver and vice versa. The speaker role is interchangeable and not fixed.

5 *Feedback*: The speaker has access to the message and can monitor its content. This allows us to monitor and correct errors or slips in spoken language and, as Hockett pointed out (1960, p. 6), this kind of cognitive access forms the basis for the internalization of language as verbal thought.

6 *Specialization*: The energy expended in producing the message does not alter the meaning of the message. Whether we whisper or shout the utterance, the meaning of the words remains the same, although we can change the emphasis by altering the vocal energy, or indeed communicate non-literal meaning using a change in tone (e.g. sarcasm).

7 *Semanticity*: Sounds within speech refer to objects and entities in the world: words have meaning. There is an element of semanticity in some animal calls. For example, vervet monkeys use a system of predator alarm calls with distinct calls for snakes,

eagles and leopards (Seyfarth et al., 1980). The calls are more likely to be made in the presence of other monkeys and, in particular, in the presence of kin (Cheney & Seyfarth, 2005). This use of specific calls by animals is referred to as **functional reference** and the information contained in the signal allows the signaller's conspecifics (members of same species) to react appropriately. However, there is little flexibility in their use.

> **Functional reference** refers to the use by animals of a specific call to stand for a specific object or threat.

8 *Arbitrariness*: The relationship between the spoken word and its referent in the world is arbitrary. Apart from a small number of onomatopoeic words (e.g. *hiss*, *buzz*, *cuckoo* and *murmur*), the form of the signal (the way it sounds) does not relate to its meaning directly.

9 *Discreteness*: The speech signal is composed of discrete units. Vocabulary is built up from smaller meaning units and the meaning units in turn are built from the basic sounds of the language. These sounds are perceived categorically. Change the /p/ sound in 'pin' to a /b/ sound and you have a different word with its own meaning, 'bin' (we follow the convention here of representing sounds by a letter between forward slashes, e.g. /p/).

10 *Displacement*: We can use language to refer to things that are displaced from the present situation, either in time or space. By contrast, animal systems tend to be tied to the current context. In the vervet monkeys' signalling system mentioned above, the issuing of an alarm call is triggered by the presence of a perceived threat but the calls cannot be used flexibly outside of that context. There is no way to communicate 'I saw that eagle earlier' or 'Bob says he saw a leopard' (see Harley, 2010).

11 *Productivity*: Language allows us to create novel utterances. This aspect of language is a fundamental distinguishing feature of human language and its basis lies in syntax (Chomsky, 1986). Productivity is also referred to as openness or generativity. From the earliest stages of language development, speech is characterized by novelty. We do not just repeat back speech we have heard; we say things in new ways. Likewise, we can understand novel sentences that we will not have encountered before.

12 *Cultural transmission*: A language is learned through interaction with more experienced users of the language within a verbal community.

13 *Duality (of patterning)*: Meaningful elements are created by combining a small set of meaningless units. For example, the 40 or so sounds of the English language are as and of themselves meaningless; however, they combine in meaningful ways to allow us to utter all the words in the English language.

14 *Prevarication*: Language can be used to deceive and lie. Furthermore, the messages we create may lack an obvious meaning.

15 *Reflexiveness*: We can use language to communicate about language. The sentences you are currently reading are an example of this property of language. Animal systems lack this feature; as Harley (2008) comments, 'bees will never dance a book about the psychology of the bee dance' (p. 57).

16 *Learnability*: A language can be learned by a speaker of another language.

These features are not independent, as Hockett noted. For example, semanticity and arbitrariness are related: words have meaning, they refer to something in the world (semanticity) and the relationship between the sound of the word and the thing it refers to is not physically direct (arbitrariness).

The design features apply to spoken language and do not apply fully to sign languages (which it is generally agreed show similar linguistic properties to spoken language) or to written language. However, they provide a useful general way of differentiating animal communication systems from language. While animal systems have some of Hockett's properties for language, the full set of design features is only found in language.

COMPONENTS OF LANGUAGE

Language is a structured system which uses a finite set of sounds to construct words, sentences and ultimately conversations. The components of language, from the smallest to the largest parts, are phonemes, morphemes, syntax and discourse. In order to understand language, it is useful to consider these components independently; we start with the sounds of language, phonemes.

Phonemes

Phonemes are the basic sounds that make up speech within a language and the term phonology refers to the system of sounds in a language. There are about 100 basic sound units or **phones** (as listed in the International Phonetic Alphabet) that can be used to make up words. This represents the full set of available sounds; the study of these raw sounds is called **phonetics**. No one language uses all these sounds; instead, languages use a subset of phones, and languages vary considerably in the number of sounds used. The sounds within a language are called **phonemes**. These are the smallest meaningful sound units in a language. The number of phonemes within a language varies – there are 44 in (British) English, while some Polynesian languages (e.g. Hawaiian) have as few as 12, and there are over 140 in the African language Khoisan (Chierchia, 2001). Table 12.3 shows the number of phonemes in a selection of European languages.

Phones are the basic speech sounds.
Phonetics is the study of speech sounds.
Phoneme is the smallest meaningful sound unit within a language.
Allophones are phonetic variants of the same phoneme.

Some basic vowel sounds occur in all languages, but consonants can be used differently and are perceived differently. For example, in English, there is no perceived difference between the aspirated (i.e. said with a puff of air) /p/ sound in *pin* and the unaspirated /p/ in *spin*, but in Thai these are two distinct phonemes. Similarly, in English the /t/ sounds in *tea* and in *trip* are different phones, but these are treated as one phoneme; if you substitute the *t* sounds, the pronunciation may sound a little odd, but it is still the same word. Phonemes also change as a function of the surrounding sounds in words and in sentences, an effect referred to as co-articulation. Different phones that are treated as the same phoneme within a language are called **allophones**. Phonemes therefore do not correspond directly to physical sounds; rather they are 'abstract representations of the phonological units of a language, the units used to represent the forms of word in our mental lexicons' (Fromkin et al., 2003, p. 285). A phoneme is a rather subjective category that is recognized as meaningful by the speakers of a language, but is not necessarily constant as objectively measured (see Chapter 13).

Table 12.3 The number of phonemes in a selection of European languages (from Tambovtsev & Martindale, 2007).

Language	Number of phonemes
Finnish	56
German	51
Italian	49
English (British)	44
Swedish	41
Dutch	39
French	36
Norwegian	36
Greek	25
Portuguese	25
Spanish	24

Source: SKASE Journal of Theoretical Linguistics

The tendency to perceive the difference between two allophones decreases with age (e.g. Iverson et al., 2003), suggesting a critical period beyond which the adult is 'tuned' to the sounds of his or her native language. While a child can discriminate between the full set of phones, an adult generally cannot; a child who is exposed to a second language can acquire native-like ability and accent, while the parents, coming late to the language, may struggle to acquire it and never acquire a native-like accent (see Bongaerts et al. 1995; Dewaele, 2009; Singleton, 2001). For example, the difference between /l/ and /r/ sounds in English is not readily discriminated by Japanese adults, for whom /l/ and /r/ are treated as allophones (Ingram, 2007). While young children can appreciate the difference, by adulthood this ability is reduced. This reduction in discrimination with linguistic experience may serve to reduce the ambiguity in the incoming speech signal, facilitating language comprehension, an issue we return to in Chapter 13.

Phonological and **phonotactic rules** describe which sounds can go together in a given language. For example, in English, a /t/ sound does not follow a /b/ sound and [ng] can occur at the end of the word (as in 'king') but not at the beginning. These rules differ across languages so that sounds that are 'natural' and easy to produce and discriminate in one language may not be so easy for adult speakers of another language. Speech segmentation relies on knowledge of word boundaries using information about phonotactic probabilities in a language (see Chapter 13).

Phonotactic rules stipulate which combinations of sounds are 'permitted' in a language.

Changing a phoneme within a word will change the meaning of that word, for example *bat* and *pat*. This is an example of a minimal pair, so called because the words differ by just one phoneme. All spoken languages use vowels and consonants, but, as outlined above, the exact set used varies across languages. Maddieson (1984) examined a 300 language sample and found that the number of consonants varied from 6–95 sounds, with a mean of 23, while vowel sounds varied from 3–47, with a mean of 9.

Morphemes

Morphemes are the meaning units of a language. They are the building blocks of words. A single word may consist of several morphemes. The term **morphology** refers to the study of the rules in a language according to which words can be constructed. Morphology can be considered as a special case of syntax (see below).

Morphemes are the meaning units of language.
Morphology is the level of linguistic analysis concerned with morphemes and their role within words.
Free morpheme is a morpheme that can stand alone as a word.
Bound morpheme is a morpheme that cannot form a word on its own, but forms a word when attached to a free morpheme.

In English, regular plural nouns are created by adding –s to the end of a word, for example, *one car* but *two cars*; *one horse*, *two horses*, etc. In these examples, there is one morpheme in the singular forms (*car*) but two in the plural (*cars*), the stem or root word (car) and the plural suffix inflection (-s). The noun *car* is an example of a **free morpheme**, as it can occur on its own, whereas the plural form –s is a **bound morpheme**, because it does not carry meaning unless it is attached to a free morpheme. Here, the –s is an example of an inflectional morpheme; it serves a grammatical function but does not change the syntactic category of the word to which it is attached (*car* is still a noun when the –s is added to make *cars*). Similarly, the verb endings –ed and –ing are inflectional morphemes. Some bound morphemes, like –ify, –ish, –able and –ment, are derivational morphemes, as they create new words with new meaning when added to a stem. They can change the grammatical category of the word. For example, the verb *develop* becomes a noun when you add –ment to give *development*. Words can be altered by adding a morpheme to the start of the word (a prefix) or to the end (a suffix) and language-specific rules govern the ways in which words can be altered. The verb *depend* becomes *dependence* when we add the suffix –ence, meaning 'condition' or 'state'; adding the prefix –in, meaning 'not', yields *independence*, and so on.

Such alterations apply only to content words (e.g. nouns, adjectives and verbs); these are open class words which can be altered or invented as usage changes. In some languages inflections on content words can be particularly informative. In Hungarian, for example, the morpheme at the end of the word indicates the word's role in the sentence, and codes whether it is a direct or indirect object. For example, consider these sentences in English and in Hungarian (example from Hoff, 2005):

The boy gave a book to the girl.

A fiú egy könyvet adott a lánynak. (The boy a book gave the girl.)

Here, 'book' is the direct object in the sentence (it is given), while girl is the indirect object (the book is given to the girl). While in English we tend to rely on word order, in most cases, to understand the role of the word in the sentence, in the Hungarian sentence above the morphemes *et* at the end of 'könyvet' and *nak* at the end of 'lánynak' give the role of the word. In this example, content words are altered to indicate the word's role in the sentence; in grammatical terms the inflections are accusative and dative case markers, respectively.

Function words provide grammatical structure that shows how content words relate to each other within a sentence.

Function words, the words that do the grammatical work of a sentence do not change (prepositions, for example; see Table 12.2) – they are a closed class of morphemes. As we will see, content words and function words are to some extent treated differently in language processing.

Semantics and the lexicon

Morphemes make up words, which in turn make up our vocabulary. Our knowledge of words and their meanings are stored in a kind of mental dictionary called the mental lexicon. The lexicon is a part of the semantic memory system (see Chapter 5). It holds our store of words and associated knowledge and links words with our general knowledge about concepts and the world. As adults, we have a store of tens of thousands of words, from which we normally have immediate access to target words as we construct a sentence. Only occasionally will we experience difficulty in calling a target word to mind, a temporary failure referred to as the tip-of-the-tongue effect (more on this later in this chapter).

It is difficult to estimate the size of an adult's vocabulary; some studies suggest that adults know about 70,000 words (e.g. Bock & Garnsey, 1998; Nagy & Anderson, 1984). Miller (1977) estimated that the average English speaking college student had a vocabulary of about 150,000 words; subsequent estimates put the figure closer to 20,000 word families, where a word family consists of a base word, and inflected and derived forms (e.g. Goulden, Nation & Read, 1990). The range of estimates reflects the difficulty in accurately measuring vocabulary and the importance of defining the unit of measurement.

Words are symbols; they are meaningful sounds and generally have a particular referent. A word might be defined as 'the smallest unit of grammar that can stand on its own as a complete utterance' (Crystal 1997, p. 440); in writing, words are generally separated by spaces. People also use other meaningful sounds, though not all are words. For example, we might use a groan to signal disagreement or a yawn to signal boredom, but these are not words. A few words are not referential, that is they have no clear referent – greetings and social conventions (e.g. saying *hello*) for example.

Word is the smallest unit of grammar that can be meaningfully produced on its own; it can consist of one or more morphemes.
Semantics is the study of meaning.

The question of what words mean and how they relate to each other raises some complex issues. **Semantics** refers to the meaning of words and morphemes and the relationship between the words we use and the objects they refer to in the world.

Syntax

We construct novel sentences when we speak; we do not generally repeat back or 'parrot' previous productions. Imagine you are telling a story to a friend when another friend joins the group and you have to start your story over. The chances are, though the meaning or gist of your story will not change, the exact sentences you use will differ. This reflects the **productivity** of human language; we do not rely on rote or stock phrases, or on memory for practised utterances. Instead we create new sentences as and when we need them. This is evident from the earliest stages of syntactic development in young children.

Productivity of language refers to the ability to generate novel utterances.
Phrase is a group of words referring to a particular idea.
Slang describes an informal pattern of speech that is considered to be 'non-standard'.

Two aspects of the language system allow us to use language productively: syntax and morphology. The term syntax describes the rules that determine the construction of **phrases** and sentences in a language. It relates to grammar but the term syntax is used more often than grammar, to differentiate it from the notion of 'prescriptive' grammar. Prescriptive grammar reflects conventions for sentence construction and is based on tradition and language prestige rather than actual language use. For example, split infinitives ('*to boldly go* where no-one has gone before') and prepositions at the end of sentences ('prepositions are not good words to end a sentence *with*') violate conventions of prescriptive grammar, but are often found in everyday speech.

Similarly, **slang** may not always meet with approval, but could nevertheless be syntactically correct: the ambiguous 'I don't know nothing' might not be considered 'good form', but the same sentence would never be uttered as 'know I nothing don't'. The study of syntax reflects descriptive grammar, that is, it reflects how language is used.

Sentences follow a hierarchical structure and are made up of two parts: a noun phrase (NP) which contains a noun, often the **subject** of the sentence, and a verb phrase (VP) which contains the verb and conveys the 'action' of the sentence. For example, in the sentence 'Sarah drank the coffee', Sarah (the subject of the sentence) is the NP and 'drank the coffee' is the VP. English, French, German and related languages use a subject-verb-object or SVO word order, that is, in a declarative (active voice) sentence the subject (or agent of the sentence) comes first, followed by the verb and then the **object** of the sentence. The order in which the words occur determines the meaning of the sentence; to use Pinker's (1994) example, 'dog bites man' is not newsworthy, but 'man bites dog' is (p. 83). The most common word orders are SOV and SVO (Greenberg, 1963) and although there are examples of the six possible types (SOV, SVO, VSO, VOS, OVS, OSV), OVS and OSV are extremely rare (Dryer, 2005; Song, 2001). Some languages have more flexible words orders; for example, Japanese is mainly SOV and Russian is SVO but both languages can use other word orders because of their use of case markings. This agent-first bias in world languages is not restricted to spoken languages. It is also found in the 'homesign' produced by deaf signers with hearing parents (Goldin et al., 1990). It also appears in second languages acquired without explicit instruction (Klein & Perdue, 1997).

Subject of a sentence is the word or words that gives what the sentence is about or performs the action.
Object of a sentence is the word or words that receives the action, or is acted on, by the subject of the sentence.

One key property of syntax underlies the productivity of sentence construction. **Recursion** refers to the repeated application of a rule and, using recursion, the same rule can be applied again and again to create a novel

Recursion refers to the ability to extend sentences infinitely by embedding phrases within sentences.

utterance. Recursion has been argued to be an essential property of human language (e.g. Chomsky, 1986). Embedded sentences make use of this property, and sentences can in principle (though not generally in practice) be extended indefinitely. For example, the English language nursery rhyme 'The house that Jack built' is an example of a cumulative rhyme using recursion:

> This is the house that Jack built.
>
> This is the malt that lay in the house that Jack built.
>
> This is the rat that ate the malt that lay in the house that Jack built.
>
> This is the cat that killed the rat that ate the malt that lay in the house that Jack built.
>
> This is the dog that worried the cat that killed the rat that ate the malt that lay in the house that Jack built . . .

Recursion would seem to be a resilient property of human language as even young children who have been deprived of language input retain the ability to use recursion (Goldin-Meadow, 1982). The extent to which recursion is uniquely human has been challenged, however. For example, songbirds have been shown to be sensitive to recursion and can classify novel patterns accordingly and reliably reject 'ungrammatical' patterns (Gentner et al., 2006).

Discourse

Discourse refers to multi-sentence speech and includes dialogue, conversation and narrative. **Pragmatics** refers to the understanding of the communicative functions of language and the conventions that govern language use.

Discourse refers to multi-sentence speech and includes dialogue, conversation and narrative. At this 'higher' level of language function, the social conventions that affect language processing become increasingly relevant and people rely on schemas (see Chapters 5 and 7) in order to process language. **Pragmatics** refers to the understanding of the communicative functions of language and the conventions that govern language use. At the level of discourse, the function of language in communicating directly and indirectly comes to the fore. A distinction is made between linguistic competence, which refers to our ability to construct sentences, and communicative competence, which refers to our ability to communicate a message effectively (Hymes, 1972). Language can be perfectly well formed, but if we fail to appreciate the social conventions governing its use, we may not communicate as we intended.

Effective discourse is based on a shared understanding between those engaging in a conversation. For example, if two people are conversing and one asks the other a question, there is an implicit agreement that the response will be related to the question. Similarly, participants in a conversation are expected to adhere to the topic of the conversation. If someone wishes to deviate from the topic or to change the subject, it is customary to signal this change of focus, by prefacing the utterance with 'by the way . . .', for example.

Conversations require turn-taking and cooperation and participants follow a set of implicit social conventions. A variety of verbal and non-verbal signals serve to regulate the conversation by indicating who speaks when and for how long. These turn-taking cues act to minimize overlap between speakers and reduce gaps or silences in conversation. Conversational turn-taking has several features (Sachs et al., 1974). One party speaks at a time; the person speaking changes. The duration of a turn is not predefined; the order of turns also varies. Transitions between turns are coordinated; overlap is minimized. These patterns hold in the absence of face-to-face information, for example, in telephone conversations (De Ruiter et al., 2006). Despite the differences in linguistic features across languages, there are striking universals in turn-taking patterns, as is explored in Box 12.1.

Box 12.1 Research Close Up: Cross language universals in conversational turn-taking

Source: Stivers, T. et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 10587–10592.

INTRODUCTION

Stivers et al. set out to examine the extent to which there are cultural differences in turn-taking in everyday conversations. They tested two opposing hypotheses. The 'universal system' hypothesis predicts little cross-linguistic variability and predicts that most languages will use a 'minimal-gap minimal-overlap' convention, as in found in English. On the other hand, the 'cultural variability' hypothesis, based on anthropological accounts, holds that turn-taking practices differ considerably across languages and cultures.

METHOD

Stivers et al. (2009) analysed video recordings of informal conversations in 10 languages from five continents. The languages varied in structural properties (e.g. word order, grammar) and were drawn from different cultures, ranging from hunter–gatherer groups to large-scale industrialized societies. All conversations were spontaneous, informal conversations, with 2–6 participants. Questions and responses were timed, coded for their form and function, and coders judged whether the responses were delayed.

RESULTS

Striking similarities emerged across the languages supporting a 'minimal-gap minimal-overlap' norm. While there was some variation across languages, the mean response time for a turn transition was very similar across languages (see Figure 12.1). Of the languages examined, Danish had the slowest response time on average (+469 milliseconds), while Japanese had the fastest (+7 milliseconds). Italian (309 milliseconds), English (236 milliseconds) and Dutch (109 milliseconds) fell between these extremes. The mean response time across languages was +208 milliseconds, and each language's mean was within 250 milliseconds of the cross-language mean: as Stivers et al. note, this is about the time it takes to say a single syllable in English. In other words, responses tended neither to overlap nor be delayed by more than half a second.

Furthermore, the factors that predicted the speed of a response were

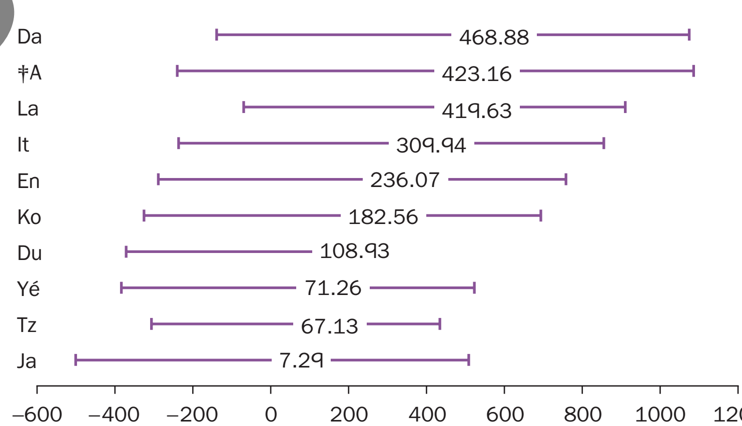

Figure 12.1 Mean time (in milliseconds, shown on the x axis) of turn transitions for 10 languages. Speakers of all languages have an average offset time within 500 milliseconds, while averages vary considerable across the languages. Languages shown along the y axis are: Da, Danish; ǂĀ, ǂĀkhoe Haillom; La, Lao; It, Italian; En, English; Ko, Korean; Du, Dutch; Yé, Yélî-Dnye; Tz, Tzeltal; Ja, Japanese.

Source: Stivers, T. et al. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS*, *106*, 10587–10592.

identical across the languages. When visible responses were made (e.g. a headshake, a nod), they occurred faster than speech in all the languages. Confirmation responses were faster than disconfirmation responses. Questions accompanied by a gaze received faster responses than those without, and answer responses (e.g. 'yes') were significantly faster than non-answer responses (e.g., 'I don't recall').

DISCUSSION

These data support strong universals in turn-taking patterns across languages, and suggest a common pattern whereby the gaps between turns, and overlaps, are minimized. While there were some differences between languages, in real terms these differences are very small. The 'long silences' associated with Nordic languages, for example, evident in the Danish response times, amount to a delay equivalent to about one syllable. The data were limited to question–response sequences, however; it would be interesting to see if similar findings apply to other types of interactions.

Grice (1957, 1975) identified a set of four conversation rules or maxims that describe effective conversations and reflect the expectations of listeners. Grice's four maxims are:

1. *Maxim of quantity*: The speaker should provide enough information in order to be understood but not too much information.
2. *Maxim of quality*: The speaker should provide accurate information.
3. *Maxim of relevance*: The speaker should provide information that is relevant to the current topic of conversation.
4. *Maxim of manner*: Ambiguity and vagueness should be avoided.

If the maxims are violated, more cognitive processing is required to determine the response, or the participants may have to backtrack or repair the conversation. Of course, violation of these rules can also form the basis of humour – hyperbole, sarcasm and irony violate the maxim of quality.

Thus far, when we consider the meaning of words we have treated this as what the word *denotes*, that is its literal meaning; a word denotes its referent. We are also able to communicate indirectly, however. If you were sitting next to an open window and someone said to you 'It is cold', the utterance might be taken as meaning 'Can you close the window?' We must also consider the connotations that words evoke. Connotation refers to the non-literal aspects of word meaning and reflects social and cultural factors that affect the literal processing of word meaning. Words can be perceived as having a positive or negative connotation (see Jay & Danks, 1977). If you hear someone say, 'Bob eats like a pig', it is of course literally true, in that the manner of eating (the movements of the jaw, and so on) is similar in many animals; however, it is likely that something else is meant. It is likely that Bob is being insulted and the connotation created by the phrase could mean that Bob is a messy eater or that he is greedy. Similarly, words can attract a positive value through use and this affects understanding of their literal meaning. The word 'natural' is much used in advertising, for example, and its use relies on the fact that people perceive 'natural' as having a value: 'natural' is equated with 'good'. The many entirely natural but life threatening infections that assail humanity do not seem to be evoked by the term, nor do 'natural disasters' spring to mind.

Having considered the constituents of language we now move on to examine how language is put together so that phonemes, morphemes and syntax become discourse. Much of what we know about this process is derived from the study of speech errors.

SPEECH ERRORS

Speech is produced at a rate of about 15 speech sounds and 2–4 words per second (Levelt, 1989) and for the most part is fluent and well formed. Slips and errors in spontaneous speech are quite rare, with some studies finding errors less than once per 500 uttered sentences (Dell et al., 1997; Garnham et al., 1981; see also Levelt, 2001). Error rates are low despite the rapid rate at which sounds must be selected from a production vocabulary of about 20,000 words (Groome et al., 1999). Hesitations and pauses in spontaneous speech are common, however.

Many of the theories of speech production outlined later in this chapter originate from analyses of speech errors. Data about errors come from several sources. The first source involves examining temporary breakdowns in the system's functioning, which occur from time to time, under normal conditions. Speech errors such as slips of the tongue and tip-of-the-tongue effects (when we cannot fully access a target word from the lexicon) are examples of this type of 'malfunction'. Second, errors can be induced in the laboratory, by having people articulate very quickly, for example. Third, the study of acquired brain injury has shown how damage to certain brain areas affects speech and language processing. **Aphasia** is the term used for acquired disorders of language, a topic we return to shortly. First, we look at hesitation and pauses in normal speech.

Aphasia is the term given to a group of speech disorders that occur following brain injury.

HESITATIONS AND PAUSES

Disfluencies such as pauses are more common than actual errors and vary with the situation and the individual. They are a natural characteristic of fluent speech. Estimates suggest that about six in 100 words are affected by disfluency (Fox Tree, 1995). These pauses can be silent or filled (common fillers include *um*, *ah* or *er*; see Maclay & Osgood, 1959, for an early account of their use). Filled pauses occur with less frequency than silent pauses (O'Connell & Kowal, 2004) and may serve to announce a delay in speech (Clark & Fox Tree, 2002). During pauses, the speaker plans the articulation of their next words (Butterworth, 1980). Goldman-Eisler (1968) found that when participants were instructed to speak about a given topic, pause duration accounted for as much as half of the total time (although the method used to arrive at this estimate has been questioned; O'Connell & Kowal, 2004). The use of pauses varies with context, task demands and from individual to individual. Schacter et al. (1991) found that pauses in academic lectures varied with academic discipline, with science lecturers producing fewer pauses than humanities lecturers. This may reflect the more precise terminology employed by scientists; if there are fewer words to choose from, word choice is facilitated. Pauses within **clauses** and sentences would seem to reflect formulation of ideas and word selection (Velmans, 2009).

Disfluency is a hesitation or disruption to the normal fluency of speech. By contrast, the term 'dysfluency' is used to refer to an abnormal disruption to fluency, such as following brain damage. The use of the prefix 'dys-' signals 'abnormal'.

Clause is a part of a sentence containing a subject and verb.

Some disfluencies may facilitate comprehension. For example, Fox Tree (2001) found that hearing an 'uh' aided listeners' recognition of subsequent words, suggesting that some disfluencies may act as cues that direct listeners' attention towards a particular word (see Box 13.1 in the next chapter). Hesitations have also been studied as potential cues to deception, as is explored in Box 12.2.

Box 12.2 Practical Application: Using verbal cues to detect a liar

There is a widespread misconception that it is easy to tell a liar from his or her demeanour, that certain behavioural cues reliably indicate when someone is lying. In fact, the average person cannot detect lies reliably and attends to the wrong cues when attempting to do so. Even people who might be expected to have had considerable exposure to deception are not necessarily effective at lie detection. Ekman and O'Sullivan (1991) examined lie detection by a sample of 509 people, including members of the US Secret Service, CIA, police and judges, as well as college students and other adults. They showed participants a videotape of 10 people who had been instructed to lie or tell the truth. Only the Secret Service personnel could detect lies at an above chance level, with average accuracy in this group of 64 per cent.

What can psychology tell us about how to detect lying? Vrij et al. (2001) note that someone who is lying, compared to someone telling the truth, tends to have a higher voice pitch, produces more hesitations and speech errors, speaks more slowly and uses fewer illustrators (e.g. hand gestures), and shows decreased hand, leg and foot movements. Furthermore, the liar's account of events produces fewer details than a true account (Vrij, 2004). These features are associated with the increased cognitive load involved in lying and are also observed when people engage in a cognitively complex or challenging task. People *perceive* lying to be associated with the following features: higher voice pitch, slow speech, more hesitations and errors, delayed responses to questions, averting eye gaze, smiling more often, and increased movements (Vrij, 2000). We therefore seem to attend to the wrong cues when attempting to detect lying, while people who attend to speech-related factors rather than behavioural cues would seem to be more effective at lie detection (Mann et al., 2004; Vrij & Mann, 2001).

Many studies of lie detection involve laboratory manipulations, with 'liars' instructed to deceive under various conditions. In real-life, lying may present differently, particularly when the stakes are high and the consequences of getting caught out in a lie are serious. Vrij and Mann (2001) examined the behaviour of a murderer who had been interviewed by the police on several occasions. The series of interviews had been videotaped. While the man initially denied that he was involved in the victim's death, the evidence suggested otherwise. He subsequently confessed to the crime and was convicted of murder. The videotaped interviews therefore allowed the researchers to examine the man's behaviour while lying (pre-confession) and telling the truth (during his confession).

While lying, the man produced longer pauses, slower speech and had more speech disturbances than when telling the truth, features that are consistent with lying carrying a heavier cognitive load. In a second part to Vrij and Mann's study, 65 police officers watched video fragments that had been selected from the interviews. The overall accuracy rate in the experiment was 64 per cent, which was significantly above chance; however, this resulted from good detection of truth rather than accurate detection of lies. In fact, accuracy of lie detection (at 57 per cent) was not above chance. The individual differences in lie detection were striking, and may reflect differences in selection of cues. Mann et al. (2004) found that police officers' accuracy was negatively correlated with reliance on stereotypical but non-diagnostic cues such as gaze aversion and fidgeting.

Such knowledge is now being applied to improve interview techniques. Vrij et al. (2010) support an 'information gathering' approach to interviewing that contrasts with the 'accusatory' approach often adopted. They note the advantages of the information gathering approach over an accusatory style: it increases the amount of detail provided by a suspect and is associated with more non-verbal and verbal cues to deceit being produced; it is less likely to elicit false confessions; and it is associated with lower interviewer confidence, which safeguards against bias. In addition, Vrij and colleagues suggest asking unanticipated questions to reduce the effects of preparation on the part of the suspect and asking questions about the precise time of events if a suspect appears to be using a scripted answer. Finally, because lying is more cognitively demanding than telling the truth, increasing the cognitive load may serve to differentiate liars from truths tellers. For example, having suspects give their account of events in reverse order increases cognitive load, as does asking event-irrelevant questions.

While speech does not provide us with a linguistic version of the 'Pinocchio's nose' with which to identify a liar, it does provide some relatively diagnostic cues that might be attended to. Such an approach would seem to be more effective than relying on stereotypical body language cues.

SLIPS OF THE TONGUE

In *The Psychopathology of Everyday Life* (1924/1938/1975), Freud treated speech errors as a particular class of **parapraxes** (action slips). While his emphasis was on supposed underlying repressed thoughts, he recognized that errors could be informative as regards language processing, asking 'whether the mechanisms of this disturbance cannot also suggest the probable laws of the formation of speech' (p. 71). So-called 'Freudian slips' are errors based on a substitution of a semantically or phonologically similar word (see Table 12.4), and, most researchers would now acknowledge, they reflect the cognitive processes underlying sentence formulation, rather than unconscious motivations or conflicts (e.g. see Norman, 1981; Reason, 1990, 2000).

Parapraxes are slips of the tongue or other actions originally thought to reflect unconscious motives.

Fromkin's (1971) analysis provided the first systematic account of error types. Fromkin showed that when errors occur they are not random; in fact they are systematic and are highly informative as to the nature of the underlying processing. The majority of speech errors are sound based errors (Fromkin, 1971, 2004) and errors tend to occur at one linguistic level (e.g. affecting phonemes or morphemes). Types of error are summarized in Table 12.4.

Table 12.4 Examples of types of speech error

Type	Description	Example(s)
Anticipation	Substitutions of a sound in anticipation of a sound that occurs later in the phrase. A full word can also be produced too early within a sentence	Cuff of coffee [cup of coffee]
Perseveration	The repetition of a sound from a previous part of the utterance	proliperation [proliferation]
Transposition/Exchange errors (also called metatheses or spoonerisms)	Transposition of two segments. Exchange errors can also affect words, where two words swap places in the sentence	You hissed all my mystery lectures [missed all my history lectures] You have tasted a whole worm [wasted a whole term]
Blend	A non-word is made based on two semantically related words	Mownly [mainly/mostly] Swinged [switched/changes]
Additions	A sound is added in	Similarily [similarly]
Deletions/omissions	A sound is omitted	Slit second [split second]
Semantic substitutions including Freudian slips	Retrieval of an incorrect but semantically related target	This room is too hot [cold]
Phonological substitutions or malapropisms including Freudian slips	A phonologically similar word is selected in error. Mixed errors, in which the target word and error share both semantic and phonological features, can also occur.	Projects [products] There's a pest in every class [pet] (this could be a deletion)

Lexical bias refers to tendency for phonological speech errors to result in real words.

Analysis of speech errors points to the importance of the phrase as a unit of production, as errors rarely jump across phrase boundaries. The vast majority of morpheme exchanges occur within clauses (Garrett, 1975). Errors preserve the consonant-vowel distinction and phonological errors are in keeping with the phonological constraints of a given language (Fromkin, 1971). Exchange errors show a **lexical bias** in that they are more likely to result in a word than a non-word (Rapp & Goldrick, 2000) – for example 'barn door' becoming 'darn bore' (Nooteboom & Quene, 2008). The frequency of lexical bias has been disputed; some researchers have argued that lexical bias is not common in spontaneous speech and is more likely to be induced experimentally (e.g. Garrett, 1980), but others argue that it also applies to natural speech (Dell & Reich, 1981). It would seem that the lexical bias effect reflects both immediate feedback between speech sounds and word forms (Dell, 1986) and monitoring of inner speech producing a real word bias (Levelt et al., 1999). In other words, non-word errors are more readily detected and repaired, while real word errors can 'slip through the net' and remain undetected before being uttered (Nooteboom, 2010).

Content words tend to exchange with content words and function words with other function words. Harley (2008) found no instances of content words and function words exchanging in his corpus of several thousand speech errors. Function words and bound morphemes (such as inflections) are generally left in place when a content word or morpheme moves, a pattern referred to as morpheme stranding. The following examples from Fromkin (1971) illustrate:

- nerve of a vergious breakdown [verge of a nervous breakdown];
- a weekend for maniacs [maniac for weekends].

Boomer and Laver (1968) found that stressed and unstressed syllables did not exchange with each other; errors were consistent with the stress pattern in the utterance. Furthermore, transpositions generally stay within the same syntactic or morphological class (Fromkin et al., 2010). These patterns of error suggest a systematic process whereby a sentence is constructed such that 'the word's skeleton or frame and its segmental content are independently generated' (Levelt, 1992, p. 10) and shows that speech production is highly rule governed (Fromkin et al., 2010).

Transposition errors, or metatheses, have been noted for centuries, and have become associated, perhaps unfairly, with the Reverend William Archibald Spooner, who was warden of New College, Oxford in the early 1900s. Many transposition errors were attributed to Spooner and this class of error has therefore become known as 'Spoonerisms'. Among the errors attributed to the Reverend Spooner (see also Table 12.4) are the following:

- you were fighting a liar in the quadrangle/lighting a fire;
- work is the curse of the drinking class/drink is the curse of the working class;
- the queer old dean/the dear old queen.

Many of these are apocryphal or exaggerated; if they were as frequently produced as some commentators have supposed it would suggest an underlying pathology (Potter, 1980).

Speech errors are generally collected from spontaneous speech, either by recording speech or having participants note their own errors in diaries, and large collections of speech errors can be accumulated in this way (e.g. Fromkin, 1971; Harley, 2008). An early example of such a corpus is provided by Meringer and Meyer's (1895) analysis based on a collection of an estimated 8,800 errors in German. Errors can also be induced experimentally, with various techniques developed to elicit errors. For example, the SLIP (Spoonerisms of Laboratory-Induced Predisposition) technique introduced by Baars and Motley (1974) aims to force transposition errors in research participants within a controlled setting. Another type of speech production error that has been induced experimentally with some success is the tip-of-the-tongue (TOT) state.

THE TIP OF THE TONGUE STATE

The **tip of the tongue** (TOT) state is a temporary inability to access a word from memory; when we experience a TOT, we generally can say whether we know the word and we may have access to some information about the word, such as its initial letter, what it sounds like, or whether it is a long or short word. In languages in which the noun has gender, that information can be available in the TOT state, showing that access to syntactic category information is preserved, although access to the specific phonological word form is unavailable (e.g. Vigliocco et al., 1997). This has implications for models of speech production, as is discussed below.

Tip-of-the-tongue (TOT) state refers to a temporary inability to access a known word.
Feeling-of-knowing is a subjective sense of knowing that we know a word, and is an example of meta-memory – our knowledge about the contents of our memories.

In the TOT state the target word is known to us, but we cannot access it; we have a **feeling-of-knowing** about the target, that is, we are aware that we know the word, yet we cannot produce it. In a review of research on the TOT state, Brown (1991) notes that it:

- is universal;
- occurs about once a week;
- increases in frequency with age;
- frequently affects recall of proper names;
- often involves an available initial letter;
- is often accompanied by other words;
- is resolved on almost half of occasions.

Brown and McNeill (1966) were the first to induce the TOT state experimentally. They read participants' definitions of rare words and asked them to produce the word to which they were referring. For example, they read the definition 'a navigational instrument used in measuring angular distances, especially the altitude of sun, moon and stars at sea'. The target word in this case was 'sextant'. The inability to access low frequency targets induced a TOT state in some participants (nine of 56 participants for this particular definition). Some of the incorrect words produced by participants show that some information is available and that lexical retrieval can involve partial activation. For example, in the case of the target *sextant*, some of the errors included *sexton*, *sextet*, *compass* and *protractor*; these are phonologically or semantically related targets. Bilinguals would seem to be more prone to the TOT experience (e.g. Gollan & Acenas, 2004), suggesting that competing lexical activations increase the likelihood of a TOT state occurring. However, the small differences in reaction times observed in the laboratory do not carry profound implications for everyday functioning of bilinguals (Baker, 2006) and should not detract from bilinguals' advantages in other domains (see Box 12.3).

Box 12.3 Research Close Up: Bilingual lexical access

Source: Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, *13*, 231–252.

INTRODUCTION

The cognitive effects of bilingualism have been studied for many years and a number of processing costs have been identified. Bilinguals would seem to be at a disadvantage compared to monolinguals on tasks involving the rapid retrieval of words from memory, though ability to access semantic information about those words is not impaired (Bialystok et al., 2009). For example, there is a bilingual disadvantage on the verbal fluency task, in which participants must name in one minute as many words as they can beginning with a given letter (e.g. name words beginning with the letter 's') or belonging to a given category (e.g. names of animals).

METHOD

Sandoval et al. (2010) examined the verbal fluency of 30 monolinguals and 30 bilinguals when naming 15 semantic categories (e.g. 'types of clothing' 'supermarket items', 'spices') and 24 double-letter categories (e.g. words beginning with 'fa'), with order of block presentation counterbalanced. The bilingual participants were English dominant English-Spanish bilinguals. Participants were instructed to name as many examples as they could think of that belonged to each category, with each trial allowing one minute for responses. In Experiment 1 responses were given in English; in a second experiment, responses were made in both languages.

RESULTS

The bilingual participants produced significantly fewer correct responses compared to the monolinguals in both the semantic and letter categories. While both groups of speakers performed better in the semantic category compared to the letter category, the bilingual disadvantage was equal for both types of stimuli. Furthermore, in the semantic task, the monolinguals produced responses of higher word frequency. Bilinguals were also slower than monolinguals to produce their first response at the beginning of a trial. Sandoval et al. suggest that interference between languages accounts for these differences. In a second experiment, responses were made in both languages. The bilinguals produced fewer responses in Spanish (the non-dominant language) than in English, and they were slower to produce their first response for a category in Spanish compared to English. Comparing performance in the two languages, Sandoval et al. found that bilinguals produced more cross-language intrusion errors when using their non-dominant language, while very few intrusions occurred when they spoke in the dominant language.

DISCUSSION

These data suggest that between-language interference in bilinguals creates processing costs that negatively affect verbal fluency, at least in the context of such experimental tasks. Does that mean that bilingualism is disadvantageous in terms of cognitive performance? The balance of evidence would suggest that the advantages of bilingualism far outweigh any potential disadvantages. In addition to the cultural, social and communicative advantages associated with bilingualism and multilingualism, cognitive advantages include greater flexibility and creativity in thinking (Baker & Prys Jones, 1998) and better executive function (Bialystok et al., 2008).

Several of the most influential theories of speech production have been based on analysis of speech errors. Here, we look at three models of speech production: Garrett's and Dell's accounts, which are based on speech error data and Levelt's account which takes a different approach.

THEORIES OF SPEECH PRODUCTION

It is generally agreed that there are a number of stages to speech production (e.g. Levelt, 1989). The first, conceptualization, is a poorly understood process by which a thought forms and is prepared to be conveyed through language. The processes by which an abstract thought or idea becomes a verbal thought remain elusive (think, for example, of the stages before you say to yourself 'I wonder if I turned the oven off before leaving the house'). The second stage involves the formulation of a linguistic plan. The concept or proposition must be translated so that the thought becomes language and the sentence that we want to output is planned. This process of translating from concept to language also remains a mysterious process, and if the goal of such theories is, as Clark and Clark (1977, p. 10) suggested, 'to discover how speakers turn ideas into words' we are arguably no closer to the holy grail. Levelt (e.g. 1989) considers the formulation stage as comprising two sub-stages. During the lexicalization sub-stage, the words are selected from the mental lexicon. In order for this to occur the concept must connect with the abstract word form or **lemma**. The lemma contains semantic and syntactic information about the target word but does not yet specify its phonological form. Formulation also involves syntactic planning; during this sub-stage the order in which the selected words will be output is decided. The third stage involves articulation of the plan. During this stage the sounds for the word are accessed, the **lexeme** is specified giving the full phonological form of the word, and the motor program for speech output is planned and articulated. In a final fourth stage, the output is monitored so that corrections can be made if necessary.

Lemma is an abstract word form that contains syntactic and semantic information about the word.
Lexeme is the basic lexical unit that gives the word's morpho-phonological properties.

The theoretical approaches to understanding speech production have much in common. It is recognized that clauses seem to be an important structure for speech planning and that processing proceeds from the abstract concept to syntactical processing, to precise phonological patterns. The various models of speech production differ in terms of the emphasis placed on these various components and in the extent to which they consider the processing stages to involve serial, parallel or interactive processing, that is, whether they favour a modular or interactive view of speech processing. Modular or serial theories posit a series of non-interacting stages, with different types of processing being completed at each stage (e.g. Fromkin, 1973; Garrett, 1980, 1982; Levelt et al., 1991a). On the other hand, interactive or parallel theories (e.g. Dell, 1986; MacKay, 1987) propose a less constrained account, with multiple sources of information operating to influence speech output. The debate over which type of theory provides the more accurate account continues.

MODULAR THEORIES OF SPEECH PRODUCTION

Garrett's model

Serial or modular theories propose that speech production progresses through a series of stages or levels, with different types of processing being completed at each level. According to Garrett's hierarchical model (e.g. 1980, 1982, 1992), speech is produced via a series of stages, proceeding in a top-down manner (see Chapter 2) so that processing

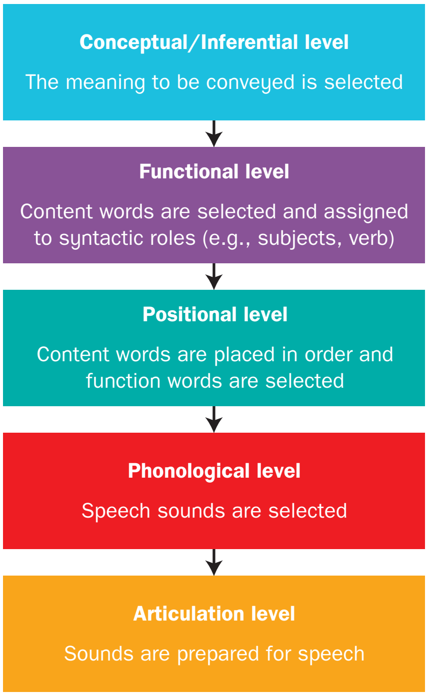

Figure 12.2 The stages of speech production proposed by Garrett (e.g. 1975). At the inferential level, we conceptualize the message that we want to express. At the functional level, the syntactic and semantic framework of the sentence is constructed. At the phonological level the sounds for the content words are acquired.

at lower levels does not influence that at higher levels (see Figure 12.2). Garrett developed his model to address patterns of errors in speech production (see also Dell, 1986, described below).

Figure 12.2 shows the various levels proposed by Garrett. At the inferential level, we conceptualize the message that we want to express. As noted previously, little is known about processing at this level or how the initial leap from thought to language-like representation might occur, particularly given that there may be a number of ways to express a thought. At the functional level, the syntactic and semantic framework of the sentence is constructed. At this stage, word exchanges would be predicted because the structure is present but the selected words, while activated, have not yet been allocated to their places within that structure or framework. Because the syntactic category of the word has already been determined, this model predicts that errors will not cross syntactic category (e.g. that nouns will swap with nouns but not nouns with verbs; verbs may swap with other verbs, though in reality verb exchanges are relatively rare). At the positional level, the words are allocated to positions within the syntactic frame. The function words are in place at this stage and so, where errors occur, bound morphemes tend to remain in their correct place (e.g. verb endings such as –ed and –ing). At the phonological level the sounds for the content words are acquired and sound errors can occur at this stage. In this model, the lexical bias effect described earlier in this chapter occurs during a later monitoring or editing stage where non-word errors are detected – word errors are more likely to 'slip through' undetected. The interactive models provide a rather different account of lexical bias, as we will see.

Evaluation

Garrett's model suggests that content and function words are treated differently and this is supported by the data on errors. One might argue that the relative sparing of function words reflects their higher frequency in language use (e.g. see Stemberger, 1985); however, bound morphemes have a lower frequency of use and yet are treated like function words in that they are retained in the syntactic frame of the sentence.

Non-plan internal errors occur when the intrusion is external to the planned content of the utterance.

Garrett's model provides a good account of the speech error data. However, Garrett's stages operate independently of each other and therefore this model does not predict errors that occur 'across levels'. For example, a type of error called a **non-plan internal error** occurs when concepts from the message level intrude when articulating the sentence, specifying the words at the phonological level. For example, someone intending to say 'let's get a coffee' while standing outside the library might say 'let's get a book'. Some 'Freudian slips' fall into this category, as a suppressed thought might interfere with current output; by Freud's account this was always the case: 'A suppression of a previous intention to say something is the indispensable condition for the occurrence of a slip of the tongue' (Freud, 1922, p. 52). Such errors have led to more interactive accounts of the stages of processing, such as that of Dell and colleagues outlined below. Before we look at interactive models, however, another influential modular account is considered, Levelt's model.

Levelt's model

Levelt and colleagues (e.g. Levelt, 1989; Bock & Levelt, 1994; Levelt et al., 1999) have presented a number of computational models of speech production leading to the sequential system called Weaver++ (Weaver stands for Word-form Encoding by Activation and Verification). This model focuses on the production of single words rather than the construction and output of whole sentences – it considers, for example, how we access (and say) the word 'cat' when we see a picture of a cat. A series of stages follow sequentially, from conceptualization to articulation (see Figure 12.3). Levelt's theory is based mainly on latency data (e.g. reaction times to picture naming) rather than error patterns, in contrast to Dell's and Garrett's models.

The theory focuses on the lexical access aspects of speech production. In Weaver++, the first two stages of processing involve lexical selection. Three stages of form encoding follow, before articulation occurs. These two systems, for lexical selection and for form encoding, would seem to involve quite different processes and involve different areas of the brain (Levelt, 2001).

The first stage is conceptual preparation, which Levelt et al. (1999) define as 'the process leading up to the activation of a lexical concept' (p. 3). The second stage involves lexical selection: a lemma or abstract word is retrieved from the mental lexicon and its syntactic category is activated. A number of words might be primed based on meaning, with selection dependent on relative activation so that the more appropriate selection occurs. For example, if I am shown a picture of a horse and asked to name it, the concepts *horse* and *animal* might be activated. These concepts activate the corresponding lexical items in the lexicon. This is an abstract word or lemma, which is 'essentially the lexical item's syntactic description' (Levelt, 2001, p. 13464).

Levelt et al.'s third stage involves morphological encoding. Once the lemma is selected, processing proceeds from the conceptual/syntactic domain to the phonological/articulatory domain; Levelt et al. recognize this as a crucial change. At this point, a TOT state can be produced – a lemma has been activated but the specific phonological form (lexeme) is not yet available. Because the lemma is a syntactic word, information about syntax is available, while the sound of the word is not yet accessed. This predicts the finding that in a TOT state a speaker has access to syntactic information, although they cannot produce the word. In many languages nouns have grammatical gender; this is unrelated to word meaning, that is, grammatical gender is a linguistic property unrelated to the conceptual properties of the referent. For example, in French *mouton* (sheep) is a masculine noun while *chèvre* (goat) is feminine. The word for 'milk' is masculine in French and Italian, and feminine in German, Dutch and Spanish. In languages with grammatical gender, information relating to noun gender can be activated in the TOT state, when the word

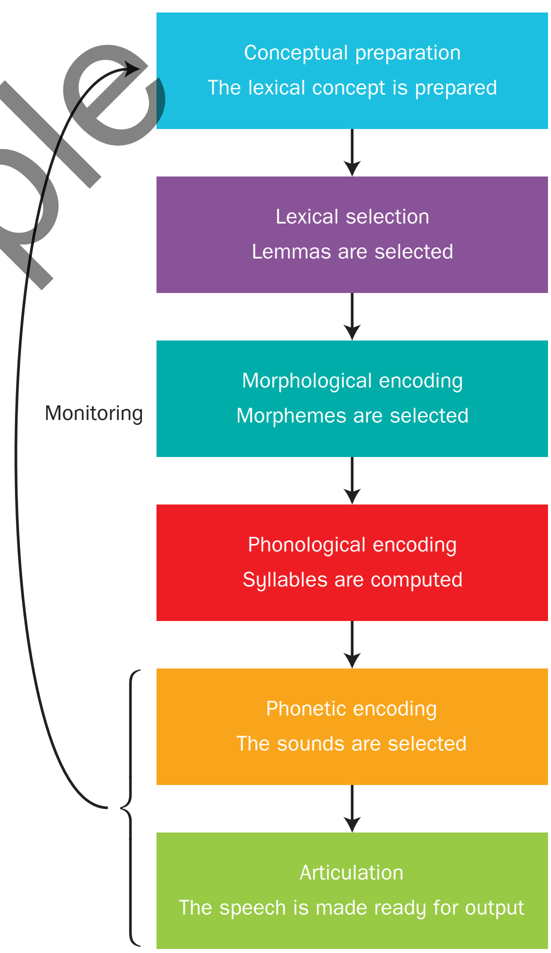

Figure 12.3 Levelt's model of speech production. In parallel to the processing stages, output monitoring allows the speaker to detect and correct errors.

Source: Adapted from Levelt, W. J. M., Roelofs, A. P. A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–37.

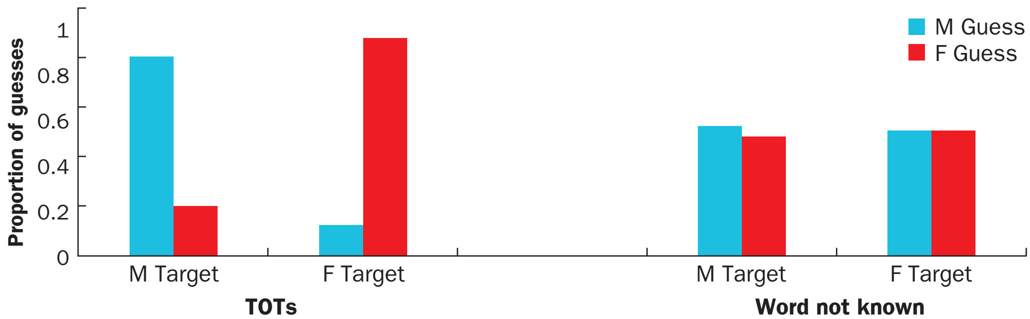

Figure 12.4 **Responses to masculine and feminine target words for TOT words and words that could not be retrieved.** Distribution of masculine (M guess) and feminine (F guess) responses for masculine (M target) and feminine (F Target) words, for TOTs compared to words that were not retrieved (here labelled 'word not known'), show access to knowledge of word gender for TOT words.

Source: Data from Vigliocco et al. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, *8*, 314–317.

itself is not accessible. Vigliocco et al. (1997) demonstrated this in Italian. Their participants were presented with definitions and asked to provide the corresponding word. Whenever a participant was unable to provide the word, they were asked to guess the gender of the noun (masculine/feminine), guess the number of syllables, give as many letters in the word as possible and state their position in the word, and report any other word that came to mind. Participants were subsequently shown the target word and asked whether it was the word they were thinking of. Vigliocco et al. found that noun gender was correctly reported 84 per cent of the time when participants were experiencing a TOT state (see Figure 12.4). By contrast, when participants could not produce the word and later could not affirm that the provided word was the target, performance was at chance level (53 per cent).

However, sometimes Italian speakers can access phonological information without accessing syntactic information (e.g. Miozzo & Caramazza, 1997), which does not support the notion of two separate stages for syntactic and phonological information – the distinction between semantic and phonological information remains. Processing at this stage involves three types of information; the word's morphology, its stress patterns, and the segments that make up the word are activated.

At stage four the syllables that make up the word are computed. Stage five performs phonetic encoding: the actual speech sounds activate at this stage. Levelt et al. (1999) posit the use of a syllabary, with phonological information allowing word articulation derived from the retrieval of syllables. These are 'highly overlearned gestural patterns, which need not be recomputed time and again. Rather they are ready made in the speaker's syllabary' (p. 5).

The sixth and final stage is articulation. The phonological information is transferred to a motor plan and executed by the articulatory system and speech musculature. The stages are presented in Figure 12.3.

The model considers the role of self-monitoring at multiple levels throughout the processing stages (Figure 12.3). This provides a mechanism for the detection of errors that allows us to repair speech and involves the cognitive mechanisms involved in speech comprehension (see Chapter 13). The precise means by which the mechanism operates and the attentional systems governing it are not elaborated.

Evaluation

Levelt's account shows how a series of specialized modules contribute to the process of speech production. It accounts for much data on speech production, including some patterns in bilingual speech. It also shows how speech might be monitored so that errors can be corrected. However, as a modular account, feedback between levels is limited. The retrieval of the word form occurs only after the lemma has been selected; there is no feedback from word form to lemma levels. However, some types of speech errors suggest that feedback does occur. Sometimes, the target word and the error share both form and meaning information, for example, saying 'rat' when you meant 'cat' (Treiman et al., 2003). This suggests that there is interference from lower to higher levels; Levelt et al. explain such errors as resulting from a failure in the monitoring processes. However, interactive models account for these and some other types of speech errors more successfully. We now turn to one influential account of an interactive type.

INTERACTIVE THEORIES OF SPEECH PRODUCTION

Dell's model

The final model of speech production that we will consider is Dell's cascaded or spreading activation account (Dell, 1986, 1995; Dell & O'Seaghdha, 1991; Dell et al., 1997), which is based on connectionist principles (see Chapter 1). This model uses the concept of spreading activation in a lexical network to show how competing activation across different levels might predict speech errors. In this model, activation from one level can affect processing at other levels, that is, processing is interactive. Processing is also parallel such that information can be processed at the different levels at the same time. These features, interactive and parallel processing, differ from the features of the serial models such as Levelt's.

There are four levels in Dell's model with processes corresponding to those described for Garrett's model above: that is, a semantic level, a syntactic level, a morphological (word) level and a phonological (sound) level. The semantic level is where we conceptualize what it is we want to say; at the syntactic level the structure of the sentence is devised; at the morphological level the morphemes that make up the target words are selected and at the phonological level the sounds that make up those words are activated.

Figure 12.5 illustrates the levels and connections. The connections between the layers allow bidirectional spreading of activation. That is, a word unit can activate the phonological units at the layer below (top-down spreading) and the semantic units at the layer above (bottom-up spreading; Dell & O'Seaghdha, 1991). According to this model, lexical access involves six steps:

1. The semantic units are activated by an external source (e.g. information from vision in a picture naming task, when you see a picture of the concept to be lexicalized or translated into words).
2. Activation spreads through the network.
3. The word unit with the highest level of activation is selected and linked to the syntactic frame for the sentence, in the appropriate slot. Once the word has been placed in the frame, its activation reduces to zero.
4. When the time is right, based on the slot in the syntactic frame the word is assigned to, the phonological information activates. If a single word is to be produced (e.g. in a picture naming task), selection of the word triggers the phonological information.
5. Activation continues to spread, but phonological units linked to the selected word become more highly activated.

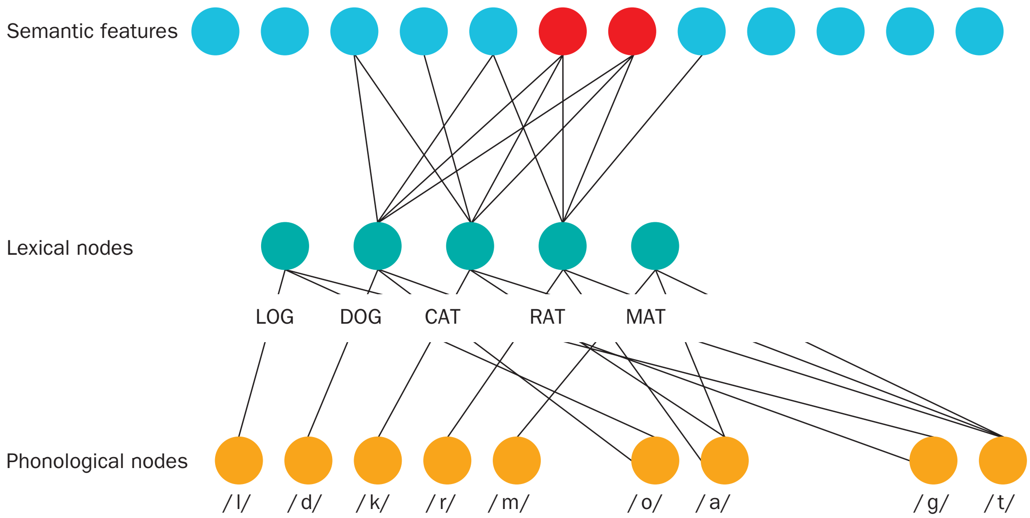

Figure 12.5 **Dell's spreading activation model.** The nodes in the top layer of Figure 12.5 represent the semantic features of the words. The words cat, rat and dog share semantic features and shared nodes are highlighted in red. The nodes in the middle layer represent lemmas or words. The nodes in the bottom layer represent the sounds that make up the start of the word, the vowel in the middle of the word, and the sound at the end of the word. Activation spreads throughout the network and the word that receives the most activation is the one that is selected (see also Levelt, 1999).

Source: Adapted from Dell, G. S., & O'Seaghdha, P. G. (1991) Mediated and convergent lexical priming in language production: A comment on Levelt et al. *Psychological Review*, *98*, 604–614.

6 The most active phonological units are selected, and these are linked to slots in a phonological frame for the word; this allows the correct phoneme to drop into the correct 'slot' in the word so that the sounds are output in the correct order (see Dell & O'Seaghdha, 1991, pp. 605–606).

During the planning stage, the various words selected for the sentence become active; activation drops off once the word is placed in the sentence. This is an interactive rather than sequential account and feedback can occur from later to earlier levels such that phonological level activation can inform processing at earlier stages. Like Levelt's model, Dell includes a monitoring process to account for self-corrections and repairs. Errors occur when activation for a non-target overrides that of the target morpheme, phoneme or word. So word substitutions occur because a semantically related, but incorrect, choice achieves a higher activation than the target word. The lexical bias effect is accounted for by a backward activation process. Because words have lexical entries, they are represented by nodes in the network; non-words do not have associated nodes, but may have activation associated with phonotactic regularities. Words should therefore attain stronger activation and be more likely to be produced in error.

Similarly, the model accounts well for exchange errors (see Table 12.2 above). Exchanges of whole words (e.g. 'the sky in the sun') occur when a word of the same category – in this example, a noun – is dropped into the wrong slot in the syntactic frame. Once it has been output, its activation drops to zero, leaving the remaining noun to take the remaining noun slot in the sentence frame, because it has high activation and has not been selected.

Evaluation

This model accounts well for many patterns of speech error and some errors produced by people with aphasia are more in keeping with a parallel model than a serial model.

For example, Blanken et al. (2002) reported a mixing of word selection and word form access, in a German patient with aphasia, that supports an interactive account. Similarly, mixed errors (see above), in which the target word and the error share both form and meaning information, suggest that feedback does occur.

The spreading activation model deals with data from speech errors rather well. It also contributed to our understanding of sentence production rather than focusing on single word production alone, as is the case in Levelt's model. However, the model does not address the semantic level in any detail, focusing instead on the construction of syntactic, morphological, and phonological representations (Dell, 1986).

As yet there is no resolution to the debate between modular and interactive accounts of speech production, and a complete model may need to consider both modular and interactive aspects of the system. Dell and O'Seaghdha (1991) suggested that Levelt et al.'s (1991) data might be reconciled with spreading activation accounts by a 'characterization of the language production system as globally modular but locally interactive' (p. 604). The degree of informational encapsulation and interaction between components remains to be established in future research.

NEUROSCIENCE OF LANGUAGE PRODUCTION

Language involves a number of interacting brain areas, with many of the key language areas located within the left cerebral hemisphere in the majority of people. Language involves a number of cognitive processes, interacting with systems for attention, memory, perception and motor function. Sociocultural knowledge informs the ways in which we use language with others. Therefore many areas of the brain must contribute to language processing. **Neurolinguistics** is the study of the relationship between brain areas and language functioning.

Neurolinguistics is the study of the relationship of brain function to language processing.

LATERALIZATION OF FUNCTION

Sensory information coming into one side of the body is processed on the contralateral (opposite) side of the brain. Similarly, fine motor movements such as hand movements are controlled by the contralateral cortical hemisphere. Information presented to the left visual field is processed in the right hemisphere, while information presented to the right visual field goes to the left hemisphere. The right side of the brain controls the left hand, and the left side of the brain controls the right hand. Information presented to each ear is processed in both hemispheres, but precedence is given to the contralateral side; stimuli presented to the right ear are processed in the left hemisphere.

Scan to watch a video on split brain experiments

Different functions are associated with the left and right cortical hemispheres. Language is largely a left hemisphere function while the right hemisphere is specialized for functions related to spatial/holistic processing (e.g. see Springer & Deutsch, 1981). When a cognitive function is *lateralized*, one cortical hemisphere is dominant for that function; for example, in most people the left hemisphere is dominant for language and the right hemisphere is dominant for face recognition. This **lateralization of function** is particularly apparent when we consider the effects on cognitive processing of a set of conditions that gives rise to the split-brain phenomenon. When the band of fibres connecting the two hemispheres, the corpus callosum, is severed, we can isolate the functions of the two hemispheres. In rare cases, these fibres may be severed surgically, to treat epilepsy for example. The corpus callosum may have failed to develop due to a developmental

Lateralization of function refers to the asymmetric representation of cognitive function in the cerebral hemispheres of humans and higher primates.

condition. In such cases, the difference in the hemispheres' functions becomes more visible. As Sperry (a pioneering researcher in this field, and Nobel prize winner in 1981) (1974) noted (p. 7):

> Each hemisphere . . . has its own . . . private sensations, perceptions, thoughts and ideas all of which are cut off from the corresponding experiences in the opposite hemisphere. . . . In many respects each disconnected hemisphere appears to have a separate 'mind of its own'.

The split brain is explored further in Box 12.4.

Box 12.4 When Things Go Wrong: The split brain

We are unaware of the division of labour between the left and right hemispheres because the two hemispheres of our brains communicate so effectively via a number of connecting bands of fibres between the hemispheres, called commissures. The largest of these is the corpus callosum. This band of over 200 million fibres is surgically severed in a *commissurotomy*, a rare surgical procedure which is performed in order to reduce the effects of a type of intractable epilepsy that is unresponsive to drug therapy.

After the procedure the 'split-brain patient' behaves surprisingly normally, considering such a radical operation has been performed. However, on careful testing, it is apparent that the left and right hemispheres no longer communicate and are effectively working independently. The left hemisphere is dominant for language, in most people. If an object is placed in the right hand of a (blindfolded) split brain patient, he or she can name the object, as the information is relayed to the left hemisphere and it can make contact with the speech centre. However, if the object is placed in the left hand, the patient cannot name it. The patient can, however, pick a matching object from an array of objects, using the same hand. A picture that is presented to the right visual field can be named; a picture presented to the left visual field cannot, although, again, the object can be matched given an array of choices. Interestingly, when information is presented to the right hemisphere and cannot be named, the person reports not seeing it, suggesting a close alliance between language and subjective experience and consciousness (e.g. Cooney & Gazzaniga, 2003 ; Sperry & Gazzaniga, 1967). However, the patient can select a related picture, using the left hand, but, unaware of what the right brain saw, he or she will often invent a reason for the selection (see Figure 12.6).

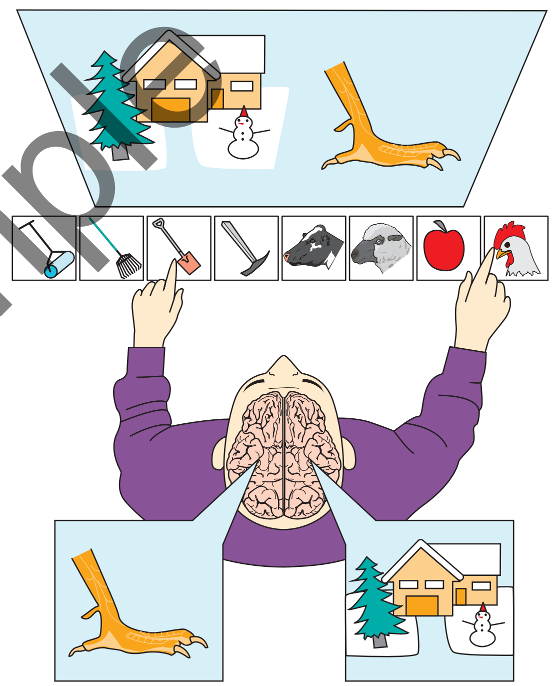

Figure 12.6 Demonstration of cognition in the split brain. A picture of a chicken claw is presented briefly to the right visual field, and goes to the (speaking) left hemisphere, while a snow scene is shown to the (non-speaking) right hemisphere. Asked to point out what he saw from a set of pictures, the patient's left hand points to a snow shovel while his right hand points to a chicken. When asked why he picked those particular pictures, the patient said 'Oh, that's simple. The chicken claw goes with the chicken, and you need a shovel to clean out the chicken shed.'

Source: Gazzaniga, M. S. et al. (Eds.) (1998). *Cognitive Neuroscience: The Biology of the Mind*. W.W. Norton & Co.

THE LEFT HEMISPHERE AND LANGUAGE

In the majority of people, speech is lateralized in the left hemisphere of the brain, and the left hemisphere is dominant for the majority of language functions. Rasmussen and Milner (1977), using the Wada test (a pre-surgical test of hemispheric dominance involving the selective anaesthesizing of the left and right hemispheres), found that 96 per cent of their patients who were right handed and over 70 per cent of those who were left handed had language lateralized in the left hemisphere (see also Kemp et al., 2008). Of the right handers, 4 per cent were right hemisphere dominant. Of the left handers, 15 per cent had bilateral representation of language, while 15 per cent were right hemisphere dominant. This dominance of the left hemisphere for language is evident in data from a number of sources: from studies of functional asymmetries in the typical population, from testing of split brain patients, and from the patterns of deficit seen in acquired language disorders such as aphasia. However, the degree of lateralization can vary in the typical population; there would seem to be, for example, differences in laterality between men and women (e.g. Shaywitz et al., 1995). Further consideration of differences in men's and women's language use is provided in Box 12.5.

Box 12.5 Research Close Up: Do men and women use language differently?

Sources: Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*, 581–592; Leaper, C. & Ayres, M. (2007). A meta-analytic review of moderators of gender differences in adults' talkativeness, affiliative, and assertive speech. *Personality and Social Psychology Review*, *11*, 328–363.

INTRODUCTION

An extremely lucrative section of the self-help industry has sprung up around the notion that men and women communicate in radically different ways, one bestselling book even relying on the metaphor that men and women 'come from different planets'. This focus on differences supports stereotypes of gender differences and ignores evidence to the contrary.

Differences in language and communication, when they are found, are surprisingly small, and the distribution of men and women on such measures mostly overlap, as two meta-analytic studies demonstrate.

METHOD AND RESULTS

Meta-analysis is a statistical technique that allows comparisons to be made across independent studies that can be combined on the basis that they address similar research questions: the effectiveness of a meta-analysis depends to a large degree on the criteria used to include studies in the analysis. Hyde (2005) reported a set of meta-analyses of research studies reporting gender differences across six categories, assessing cognitive variables such as abilities, verbal or non-verbal communication, personality variables, measures of psychological wellbeing, motor behaviours, and miscellaneous constructs such as moral reasoning. Hyde found that 78 per cent of the studies reporting gender differences showed small effect sizes (Cohen's d (a measure of the difference between two means) range of 0.00 to 0.35), with a median effect size of 0.21 (Fiske, 2010). Strong gender differences were supported only for motor skills such as throwing speed and distance, on some measures of sexuality (such as reported incidence of masturbation and attitude to sex outside of a committed relationship) and physical aggression. Measures of vocabulary, reading and speech production showed small or negligible effects, and any stronger effects are likely to be context dependent. For example, one of the stronger (though still small) effects among the communication variables suggested that men are more likely than women to interrupt a conversation (d = 0.33). An effect like this is likely to be influenced by other factors such as personality, topic of conversation and factors concerning the other interlocutors. Yet, the stereotype that men interrupt and women don't prevails.

Leaper and Ayres (2007) conducted meta-analyses of studies examining gender differences in adults' talkativeness, affiliative speech, and assertive speech. While there were statistically significant effects for all three language constructs, the average effect sizes were negligible. Contrary to the stereotype, men were found to be more talkative than women (d = 0.14) across the studies examined. Consistent with the stereotype, men used more assertive speech (d = 0.09), while women used more affiliative speech (d = 0.12). However, in all cases the effect seizes are so small that it would not be prudent to conclude that there are differences, particularly when there is inconsistency across studies examining these constructs. Studies of these language constructs in children have also produced very small effects (e.g. Leaper & Smith, 2004: talkativeness, d = 0.11 in favour of girls; assertive speech, d = 0.11 in favour of boys; affiliative speech, d = 0.26, in favour of girls).

DISCUSSION

To conclude that there are robust sex differences in speech production, we would expect to see repeated demonstrations of the difference across a variety of studies and contexts and always in the same direction. We would also expect to see consistent strong effect sizes, as occurs when we measure throwing distance, for example. This is quite simply not the case, and it would seem that gender differences in language ability have been greatly exaggerated. Wallentin's (2009) comprehensive review of sex differences in verbal abilities and language cortex concluded that there are no sex differences in language proficiency and highlighted the inconsistent findings relating to differences in language-related cortical areas. Wallentin also notes the problem of publication bias – studies that have found statistically significant differences are more likely to be published than those showing null findings. We therefore do not know how many unreported studies finding no gender differences may be out there.

EVIDENCE FROM THE TYPICAL POPULATION

Dichotic listening task is one where different stimuli are presented to each ear.

The **dichotic listening task** (see also the section on attention in Chapter 3) involves the simultaneous presentation of stimuli to the left and right ear. While auditory processing involves both contralateral (opposite side) and ipsilateral (same side) connections from ear to brain, contralateral connections are dominant; that is, while stimuli presented to the right ear are received by the auditory cortex of both cerebral hemispheres, the dominant connections are contralateral and therefore verbal stimuli presented to the right ear are predominantly processed by the left hemisphere (Kimura, 1967). Tests using the dichotic listening task have shown that there is a right-ear advantage for verbal stimuli (see Chapter 13). Participants report more words (or speech sounds) that have been presented to the right ear compared to the left (Springer & Deutsch, 1981). This left hemisphere specialization seems to be in place at quite a young age, as children as young as two years of age show a right-ear advantage for speech sounds (Hiscock, 1988) and infants under 10 months show greater left hemisphere activity when brain waves are measured during presentation of speech (Molfese & Betz, 1988). Studies suggest that the right-ear advantage may be restricted to consonant sounds (Best, 1988). Consonants and vowels may be treated differently; many non-speech vocalizations, and even the calls of apes, are vowel-like. On the other hand, the rapid changes in consonant sounds evident in human speech are complex auditory patterns and require high level sequential processing (some would argue that language is a special case of sequential processing).

Different areas within the left hemisphere process information relating to meaning and to syntax. Another way in which we can examine language in the normal brain is through measuring event-related potentials (ERPs; see Chapter 1). ERPs provide high temporal resolution (meaning that very quick changes in brain activity, with time scales of milliseconds, can be detected) and are tied to a particular event, such as a stimulus presented to a research participant. Electrodes are placed on the scalp to measure changes in electrical activity in the cortex as stimuli are presented. The changes in the brain waveform are informative as regards the nature of language processing. For example, Kutas and Hillyard (1980) compared brain waves as normal, semantically anomalous and physically anomalous sentences were presented (as written stimuli) to participants. Brain activity in response to semantically incongruous sentences differed from that seen when physically incongruous sentences were presented (see Chapter 13), supporting the notion that syntactic and semantic information are treated differently and processed in different areas of the brain (see also Hagoort & Brown, 2000; Osterhout & Holcomb, 1992).

Language has also been studied in the normal brain using a method called **transcranial magnetic stimulation** (TMS). TMS is a non-invasive way of stimulating particular cortical areas such that their functions are facilitated or inhibited. While the stimulation is short-lived and effects are largely temporary, in some circumstances TMS has been shown to alter the functioning of the brain beyond the initial period of stimulation, and therefore the method has implications in terms of therapeutic interventions. For example, Wirth et al. (2011) used TMS to enhance participants' performance on an overt picture naming task, and de Vries et al. (2010) used TMS to enhance grammar learning, or specifically the ability to detect syntactic violations. De Vries et al. showed that Broca's area plays a crucial role in grammar processing. Such studies suggest potential applications to remedial intervention in cases of language disorder after brain injury.

Transcranial magnetic stimulation is a non-invasive method of temporarily exciting or inhibiting cortical areas.

So far we have considered the left hemisphere as the site of language processing, but the right hemisphere also has a role to play, albeit a supporting role. The right hemisphere is involved in emotional aspects of speech, prosody and aspects of non-literal speech. People who have damage to the right side of the brain have difficulty in appreciating the emotional tone of an utterance (Caplan, 1987), and they have difficulty in understanding non-literal speech such as sarcasm, figurative language and indirect requests (Weylman et al., 1988), suggesting a role for the right hemisphere in processing the pragmatic aspects of an utterance. Similarly, studies of individuals with split-brain syndrome show that the right hemisphere is very limited when it comes to syntactic and phonological processing but it may be capable of other language functions, albeit not in the specialized way of the left hemisphere (e.g. Gazzaniga, 1983). While the right hemisphere is involved in the processing of tone in non-tonal languages (such as English), it is the left hemisphere that processes tone in tonal languages (such as Chinese) in which tone carries meaning (Gandour et al., 1992).

EVIDENCE FROM APHASIA

Speech production results from the processing that occurs in a number of language areas located around the Sylvian fissure of the left hemisphere, an area referred to as the peri-Sylvian language region. Damage to any of these areas can impair the ability to produce speech or writing. We can learn much about speech processing in normal cognition by examining the ways in which language is affected by brain injury.

Some of the key left-hemisphere language areas are shown in Figure 12.8; the figure does not reflect the considerable individual variability in the functional localization of

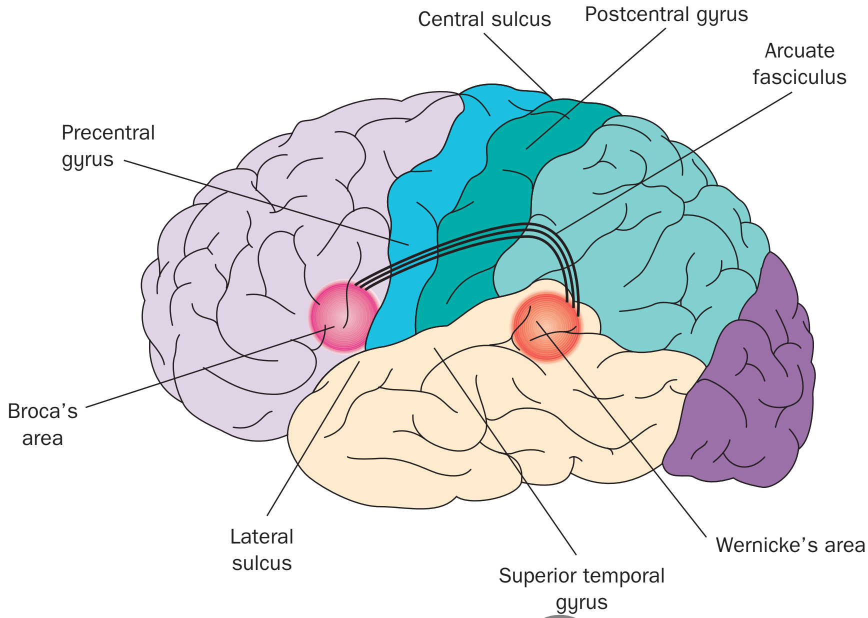

Figure 12.7 **Key left hemisphere language areas described by the Wernicke–Geschwind model.** In fact, the precise location, and role, of the language areas continue to be debated, not least because of the considerable individual variability that is evident in the functional localization of language.

Electrocortical stimulation of the surface of the cortex allows a surgeon to locate, and avoid damage to, brain regions associated with a particular cognitive function.

Wernicke-Geschwind model is a simplified model of language function used as the basis for classifying aphasia disorders.

language, however. In patients undergoing surgery to these areas, **electrocortical stimulation** while the patient is awake allows the surgeons to locate individual language areas and to reduce the risk of post-operative neurological deficits.

The **Wernicke-Geschwind model**, proposed by Karl Wernicke (1874), and later Norman Geschwind, notes a number of key areas for language (see Figure 12.7) and presents a simplified account of their role in language processing. The model proposes that we repeat a heard word by processing of the following sequence of brain areas. Following processing of the word in the auditory cortex, information about word meaning is processed in Wernicke's area and the output is sent to Broca's area via a band of connecting fibres called the arcuate fasciculus (see Figure 12.7). Broca's area prepares the speech output and a motor program for output is then articulated via the motor cortex. When we read a word out loud, a similar sequence is involved, with processing starting at the back of the brain in the primary visual cortex and continuing into Wernicke's area via the connections of the angular gyrus. While this model represents a simplification of the processing involved, it does provide a useful overview of the principal cortical brain areas for language and their functions.

Crossed aphasia refers to language dysfunction following right hemisphere damage in a right-handed individual.

Aphasia (or dysphasia as it is sometimes called) is the term used to describe a deficit in language following brain injury. It generally refers to spoken language, with the terms agraphia and alexia used specifically for deficits in writing and reading, respectively. In aphasia, the internal processing of language has broken down; it is not that the person's muscles or motor control for producing language have been damaged (as may occur in conditions such as apraxia and anarthria, for example). In a small percentage of people, damage to the right side can produce aphasia; aphasia following right hemisphere damage is called **crossed aphasia**.

Aphasic disorders can be classified according to whether they are of fluent, non-fluent or pure type. In the pure disorders a particular facet of language (e.g. the ability to repeat back sentences) is affected, while other language functions remain intact. The fluent disorders are characterized by fluent but empty speech, that is, the person produces fluent sentences, but the content of the utterances is not as they intended. The non-fluent disorders are characterized by reduced speech output, slow or effortful speech. Generally, damage to the anterior regions (near the region marked as Broca's area in Figure 12.7) creates a non-fluent type of disorder while more posterior damage (near Wernicke's) can cause a fluent type of aphasia. However, the site of damage can vary considerably from patient to patient, even with similar deficits, and younger people tend to show a non-fluent pattern of deficit regardless of the site of damage; fluent disorders are very rarely found in children, for example (e.g. Murdoch, 2009). In what follows, we will describe in a general way the main deficits associated with each category of aphasia; in fact, people with aphasia show quite a range of individual differences in performance on language tasks, and in terms of recovery of function.

Broca's aphasia

One of the first cortical areas involved in language production to be identified occupies the left inferior frontal gyrus and is known as Broca's area (see Figure 12.7). In 1861, a French doctor, Paul Broca, localized language to the left hemisphere, and attributed the production of speech to the area now named after him. (A paper by Marc Dax, dated to 1836, is now acknowledged as the first to identify the left hemisphere as the seat of language.) Broca's account was based on the aphasic disorder of a patient he encountered at the Bicêtre hospital in Paris. The man, called Leborgne, presented in his twenties with a severe reduction in speech output; over the subsequent years he gradually lost the use of his right arm and leg, an impairment confirming left hemisphere damage, as limb movement is largely controlled by the contralateral cortical hemisphere. Broca wrote of Leborgne (1861a):

Broca's aphasia is an acquired language disorder characterized by non-fluent speech, reduced speech output and problems with grammar processing.

Scan to see a patient exhibiting Broca's aphasia

> He could no longer produce but a single syllable, which he usually repeated twice in succession; regardless of the question asked him, he always responded: tan, tan, combined with varied expressive gestures. This is why, throughout the hospital, he is known only by the name Tan.

Broca called this disorder 'aphemia' (meaning 'without speech'), as he believed that the patient could understand language and therefore that the speech deficit was independent of language function itself; that it was a specific problem with voluntary speech. The similar term 'aphasia' was subsequently coined by Trousseau (Broca, 1864; cited in Dronkers et al., 2007). It is now recognized that there are some comprehension problems associated with Broca's aphasia, and these problems are particularly apparent when test sentences move beyond simple syntax (e.g. passive voice constructions). It is also now clear that it is the abstract representation of speech that is impaired in Broca's aphasia; in deaf signers with aphasia, the linguistic components of sign language are similarly affected (LeBrun & Leleux, 1986; Poizner et al., 1984).

After Leborgne's death at the age of 51 years, the area now known as Broca's area was discovered on autopsy to have been damaged by infection, which left a large abscess in that region. Broca concluded that this area of the brain was responsible for speech production. Leborgne's deficit was severe, leading some modern commentators to question whether he had a more **global aphasia**.

Global aphasia is an acquired language disorder involving extreme impairment of language function.

Non-fluent aphasia is when the patient's speech output is reduced, laboured, or absent.

Broca did not dissect the brain but preserved it and so his analysis of the damage was restricted to a surface inspection. Dronkers and colleagues (2007) were able to access Tan's preserved brain, and subject it to high resolution MRI scanning. They found substantial lesions extending into medial regions of the brain, in addition to the surface lesions that Broca reported, suggesting that global aphasia may be the more likely diagnosis.

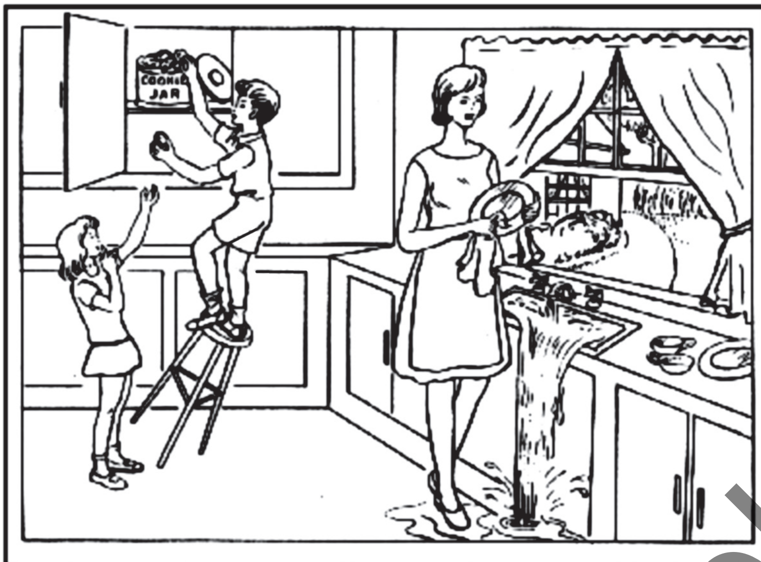

Figure 12.8 The cookie theft picture from the Boston Diagnostic Aphasia Examination. The picture shows a number of characters and actions within a familiar scene (a stereotypical kitchen scene) and elicits predictable spoken descriptions involving nouns (boy, girl, cookie, stool, silk, water) and verbs (looking, taking, spilling, falling) as well as discourse around intentions (such as to take without being seen).

Source: Cookie Theft picture. (Adapted) From the Boston Diagnostic Aphasia Examination – Third Edition, by Howard Goodglass in collaboration with E. Kaplan and B. Barresi, #11880, Austin, TX: PRO-ED. Copyright 2001 by PRO-ED, Inc. Reprinted (Adapted) with permission.

Patients with Broca's aphasia show deficits ranging from severe mutism to dysfluency or laboured speech. Broca's aphasia is one of a number of disorders that are categorized as **non-fluent**, expressive or productive aphasia. Speech output is reduced and non-fluent, but word selections tend to be meaningful. Function words (those that do the grammatical work in a sentence) rather than content words tend to be compromised. People with non-fluent aphasia are aware of their speech problems, which has implications for testing, as motivation may be an issue.

The following excerpt from Buckingham (1981) illustrates the marked dysfluency and reduced output associated with non-fluent aphasia. In this excerpt, B.L., a patient with Broca's aphasia, is trying to describe a picture from the Boston Diagnostic Aphasia Examination, called the cookie theft picture (see Figure 12.8).

The description is as follows:

B.L.: Wife is dry dishes. Water down! Oh boy! Okay Awright. Okay . . . Cookie is down . . . fall, and girl, okay, girl . . . boy . . . um . . .

Examiner: What is the boy doing?

B.L.: Cookie is . . . um . . . catch

Examiner: Who is getting the cookies?

B.L.: Girl, girl

Examiner: Who is about to fall down?

B.L.: Boy . . . fall down!

A number of features of Broca's aphasia are apparent from this excerpt. First, reduced output is apparent. This type of speech output is sometimes known as telegraphic speech, because the sentences are reduced to the most basic units required to convey meaning – the content words such as nouns and verbs. The selection of content words is correct, showing that the patient can access the words from the lexicon. The function words are by comparison relatively sparse – inflections such as verb endings, conjunctions (e.g. and, but) and prepositions (to, under) are absent.

Goodglass and Geschwind (1976) defined Broca's aphasia as a condition 'marked by effortful, distorted articulation, reduced speech output, and agrammatic syntax but sparing of auditory comprehension' (p. 237). However, as mentioned above, while comprehension of everyday conversation may be relatively intact, people with Broca's aphasia have difficulties in understanding complex syntax. When comprehension depends on processing and understanding the syntactic structure of the sentence, it fails (Cornell et al., 1993).

Broca's area is just one part of the left inferior frontal cortex; a number of areas have been identified, within that region, that contribute to various aspects of language function. For example, separate areas for semantic and phonological processing have been identified within the left inferior frontal cortex (e.g. see Gough et al., 2005). Grodzinsky and Santi (2008) provide a useful overview of the state of current knowledge around the functions of Broca's area and reiterate a key role in syntactic processing (see also de Vries et al., 2010).

Wernicke's aphasia

A few years after Broca's discovery, Carl Wernicke reported a contrasting pattern in two patients who, after brain injury, showed normal pace and intonation but jargon-like speech. Wernicke's aphasia is associated with damage further back in the brain than the region associated with Broca's aphasia, in the upper part of the left temporal lobe (and extending to the angular gyrus, and the supramarginal gyrus; this region is the inferior parietal lobule). Wernicke's aphasia is classified as a **fluent aphasia**. While speech output is fluent, it is empty, that is, it is not meaningful. This condition was sometimes referred to as neologistic jargonaphasia, the word neologism referring to the patients' tendency to produce non-words, which may reflect partial activation of phonological information (Ellis et al., 1983). These patients are generally unaware of the problem with their speech output. An example from Goodglass (1983) illustrates some of the characteristic speech patterns of the disorder, again using the cookie theft picture (see Figure 12.8):

> Well this is . . . mother is away here working her work out o'here to get her better, but when she's looking, the two boys looking in other part. One their small tile into her time here. She's working another time because she's getting, too.

Wernicke's aphasia is a fluent aphasia, characterized by fluent but meaningless output and repetition errors.
Fluent aphasia is when the patient's speech is fluent, but not meaningful.

Scan to view a patient exhibiting Wernicke's aphasia

Another excerpt from Goodglass (1993, p. 86) shows similar features. In this case, the patient was responding to being asked 'How are you today?': 'I feel very well. My hearing, writing been doing well. Things that I couldn't hear from. In other words, I used to be able to work cigarettes I didn't know how . . . Chesterfeela, for twenty years I can write it.'

People with Wernicke's aphasia are likely to produce phonemic paraphasias, that is substitution errors in which a similar sounding word (or non-word) is produced instead of the target word (e.g. 'why' for 'wine'). There are made-up words or neologisms and overall there is a striking contrast to the pattern seen in Broca's aphasia. The function words and the grammatical structures of the sentences produced are relatively intact; the problem concerns the content words. Wernicke (1874) speculated that while Broca's area was involved in motor programs for speech output, the area now known as Wernicke's was involved in processing sounds for meaning. He also speculated as to what would happen if the connections between the two areas were severed: the patient would have difficulty repeating back what was said.

Table 12.5 Summary of language deficits in aphasia and site of damage

Type	Lesion site	Effect on speech output	Other deficits
Broca's aphasia	Anterior	Non-fluent output, reduced effortful speech	Repetition Naming
Wernicke's aphasia	Posterior	Fluent but 'empty' or meaningless speech	Comprehension Repetition Naming
Conduction aphasia	Arcuate fasciculus	Fluent	Repetition Naming
Anomic aphasia	Can be anywhere in language region	Fluent but with word finding difficulty	Naming
Global aphasia	Large area of damage	Extremely limited language function	Comprehension Repetition Naming

Conduction aphasia is when the patient has a specific difficulty affecting the repetition of speech.

The arcuate fasciculus was identified as the band of fibres that connects Broca's and Wernicke's areas (in fact, it is now known to link Wernicke's and the motor/premotor frontal areas) and 'disconnection' of this band of fibres is associated with a specific deficit in repetition, a disorder known as **conduction aphasia**. Bartha and Benke (2003) outline the main characteristics of conduction aphasia: severely impaired repetition, frequent phonemic paraphasias (saying unintended syllables or words, e.g. saying 'whine' instead of 'while'), repetitive self-corrections and word-finding difficulties. Repetition deficits are a key feature; spontaneous speech is generally fluent, although paraphasic, and comprehension is close to normal. Neuroimaging studies have revealed that the neurological bases of conduction aphasia are more complex than originally thought. As Ardila (2010) notes, relatively few cases of conduction aphasia have a lesion limited to the arcuate fasciculus, and, furthermore, conduction aphasia can occur when damage is limited to the cortex, without subcortical lesions. The main symptoms of the different categories of aphasia are summarized in Table 12.5.

Anomic aphasia

Anomic aphasia is when the patient has a specific difficulty with word retrieval.

Anomia is a word finding disorder that has been likened to the TOT effect in normal speech. Relatively small lesions within the language areas can produce anomia, as can transient conditions that reduce blood supply to these areas (Obler & Gjerlow, 1999). For the individual with anomia, access to the word that he or she is searching for is denied, but the patient has not lost knowledge of the word or of its meaning. Allport and Funnell (1981, p. 405) illustrate one patient's word finding problem with the following excerpt, again using the cookie theft picture in Figure 12.8 (the square brackets shows their guess as to what the patient was trying to say):

> Well it's a . . . [kitchen] it's a place and it's a girl and a boy and they've got obviously something which is made . . . some . . . [biscuits], some . . . made . . . well . . . [the stool] it's just beginning to . . . [fall] go and be rather unpleasant . . . and . . . this is the [mother?] the woman, and she is [pouring?] putting some . . . [water] stuff. . . .

The same patient could select the correct name for an object when shown a picture and two written object names as long as the two words were not related in meaning. Therefore it is not that knowledge of words is impaired; rather the patient's ability to access the words is deficient. Allport (1983) suggests that this reflects a problem with translation between word forms and their conceptual representations. A similar pattern has been observed in developmental disorders of language such as specific language impairment (SLI). Constable et al. (1994, p. 1) reproduced the speech of a seven-year-old boy with SLI as he tried to name a set of handcuffs presented in a picture naming task:

> Key . . . oh what do you call them . . . oh yeah . . . you put . . . you put . . . with your . . . with your . . . oh . . . with your . . . when you . . . when someone's stole something . . . and . . . what do you call them . . . necklace? . . . no . . . I just don't know the word.

Evaluation

Neuropsychological cases have contributed valuable data towards cognitive accounts of speech and language production. However, we must be cautious in interpreting data from such cases. As mentioned above, there is considerable variation between patients; people with the same pattern of deficit can have damage to different areas and those with similar damage can have differing language deficits. Lesions to Broca's area can occur without Broca's aphasia (Dronkers, 1996) and a Broca's-type aphasia can follow damage to areas outside Broca's area (Caplan & Hildebrandt, 1988). Furthermore, particular types of aphasia are generally associated with a reduction in a particular behaviour (use of function words, for example) rather than a complete absence of such features (Kolk, 2007). These factors must be taken into account when considering aphasia syndromes as applied to models of normal language use.

Box 12.6 Practical Application: Supporting language expression in 'locked-in syndrome'

In the book *The Diving Bell and the Butterfly* (1997), French journalist Jean-Dominique Bauby described his life after a stroke affecting the brain stem left him with a condition known as 'locked-in syndrome'. Patients with the condition are described as 'locked in' because, although they are awake, they cannot move; they are essentially locked inside their paralysed bodies. A particularly distressing aspect of the condition is the inability to communicate with others. As Gosseries et al. (2009) note, 'testimonies from victims relate that the worst aspect of the experience was the anxious desire to move or speak while being unable to do so' (p. 192).

In 'classic' locked-in syndrome, eye blinks and up-down eye movements remain intact, allowing some patients to communicate by means of an eye blink or eye movement system. Patients with total locked-in syndrome are unable to produce any voluntary movements (Gosseries et al., 2009; Kübler et al., 2001). The ingenuity of patients in such situations is remarkable; in one case, a patient used eye blinks to communicate via Morse code (Feldman, 1971).

Jean-Dominique Bauby painstakingly dictated his 136-page memoir using eye blinks. A frequency-ordered alphabet (e.g. in English: E-T-A-O-I-N-S-R-H-L-D-C-U-M-F-P-G-W-Y-B-V-K-X-J-Q-Z) was read aloud to him – he selected the target letter by blinking his left eyelid as the letter was read. The system relies on an effective partnership between the patient and an interlocutor, who notes the letters and may become adept at predicting the patient's intended word. With practice, the system can allow the patient to communicate immediate concerns, though it does not lend itself to conversation and the patient has to rely on the interlocutor to convey his or her meaning.

Advances in augmentative and alternative communication technologies have facilitated communication for those with locked-in syndrome. Eye-controlled computer-based technology reacts to patients' eye movements, allowing them to navigate a computer system with eye movement fixations functioning much as a mouse is used to move a cursor. The computer screen can show a keyboard, from which patients select letters by eye fixation. A speech synthesizer can then allow their completed sentence to be read out loud. Menu keys also allow the user to control the environment, for example, the patient could call for assistance, browse the internet or send an email (Gosseries et al., 2009). The advantage of this system is the control and independence it affords the patient. Recent developments in brain–computer interfaces (BCI) hold further promise. BCI systems make use of brain activity, via electroencephalogram (EEG) oscillations or event-related brain potentials (ERPs), or blood flow based measures, for example, to drive external devices in real-time, essentially allowing the impaired motor system to be bypassed (Birbaumer et al., 2008). While such systems are still in their infancy and remain to be fully tested in patient populations, the results so far are promising.

WRITING

Having looked at the processes involved in the production of speech, this chapter ends with a brief overview of another form of language production: writing. The processes involved in writing are similar to those involved in speech production, but writing requires access to the orthographic (written) form of a word rather than its phonological form. Writing processes will therefore differ to some extent according to the orthographic properties of the language used (see also the section on reading in Chapter 13). Writing also differs from speaking in that when we write we have more time to think about what it is we want to express and to 'translate' it into a written form. We can take our time over the construction of sentences, whereas speaking is time pressured. We can also monitor the output more easily; we can read the sentence we have written and inspect it, and correct it if necessary. Unlike speech, writing is often a solitary activity; while a writer will have a reader in mind when writing, he or she lacks the immediate feedback that occurs during a spoken conversation. Another difference between discourse and writing is that writing makes fewer demands on memory and therefore more complex ideas can develop through writing. The act of writing ideas down can facilitate thinking and bring about deeper understanding of the subject matter (Pijlaarsdam et al., 1996).

Research examining writing focuses on the later stages of the process, including composition and revision processes, as the earlier processes such as lexical retrieval and structuring a sentence have much in common with speech processes. Writing a textbook or an essay involves a number of higher cognitive processes collectively referred to as composition. Composition is a process by which ideas are turned into symbols (Kellogg, 1999). Like speech production, composition involves the various components of language, with smaller units contributing to the overall goal of discourse-level output. As is the case for speech production, writing involves a number of stages.

THE HAYES AND FLOWER MODEL OF WRITING

Hayes and Flower (1980) proposed a cognitive model of writing that focuses on three main domains affecting the writing process. These are the task environment, long-term memory and the immediate cognitive aspects of the writing process. The task environment includes the topic of focus, the intended readership and the purpose of the writing task. The writer must have an accurate understanding of these factors in order to progress the writing task. For example, if you are writing an essay on the psychology of language, you must identify the topic and the main points, write with a reader in mind (in this case an examiner who will be grading the essay) and consider the factors that will lead to the essay receiving a good grade.

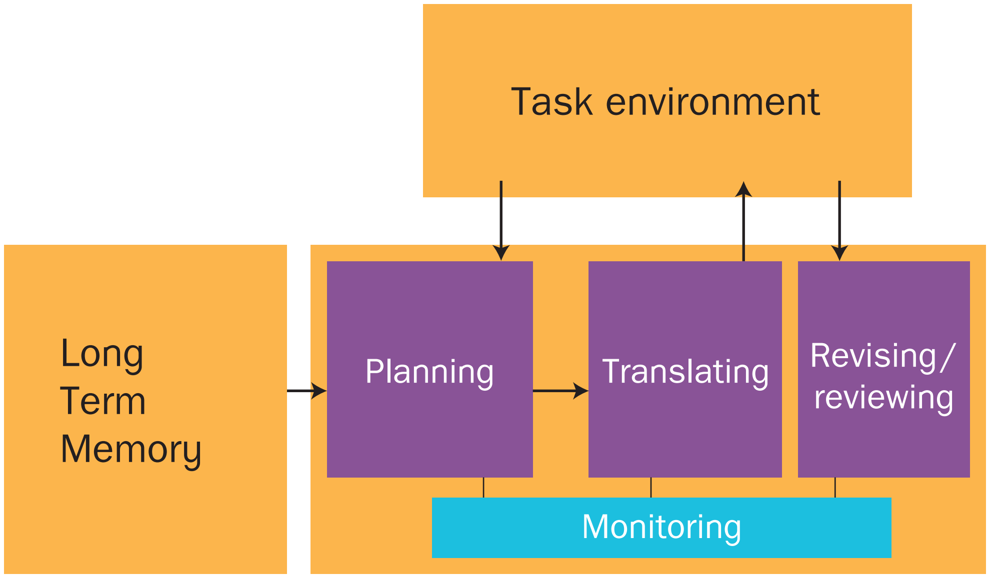

Figure 12.9 The Hayes and Flower model of writing. The model aimed to identify the key cognitive aspects of the writing process.

Source: Adapted from Hayes J. R., & Flower L. S. (1980). *Identifying the organization of writing processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

The second component of the Hayes and Flower model concerns the writer's long-term memory. The availability and accessibility of relevant information in long-term memory supports the writing process. In addition to knowledge about the subject matter, the long-term memory system stores schematic information that will shape the writer's view of readers' expectations. The third component concerns the writing process itself and the immediate cognitive demands it brings. Working memory demands (see Chapter 4) are relevant to this component.

Hayes and Flower discuss three general stages of writing: planning, translating and reviewing (see Figure 12.9). Planning includes the sub-operations of generating, organizing and goal setting. Translating converts ideas from memory into sentences on paper. Reviewing involves reading and editing. Revision is a key stage in the writing process; good writers revise more, and focus on the meaning of the text rather than the more superficial qualities of the writing (Flower et al., 1986). This model considers writing as a metacognitive act, with an executive process monitoring the key processes of planning, translating and revising (Peverly, 2006). The goal of writing is to create 'reader-based prose', as opposed to 'writer-based prose' (Peverly, 2006). Hayes's (1996) revised model of writing acknowledges the central role of working memory capacity in skilled writing.

Evaluation

The Hayes and Flower model brings together three key aspects of writing: the writing task itself, the cognitive processes involved in writing, and the writer's knowledge and long-term memory. The model moved away the previous sequential models and placed an emphasis on multiple cognitive sub-processes that allow writers to plan, revise and re-draft text. Later versions of the model also considered the role of working memory in writing. But a number of questions remain. What aspects of long-term and working memory function predict writing quality? How might memory processes support the developing writer? How does cognition differ in expert and novice writers? What role does oral language fluency play in the writing process? What role do reading skills play in the writing process? The element of time is also absent in this and other cognitive accounts of writing (see Becker, 2006): when do the various cognitive activities occur, when are certain actions initiated?

Models of writing typically have not considered the role of lower level processes in the development of writing skill, yet clearly the mechanical aspects of writing (e.g. motor skills) are an important support to the higher level cognitive processes. Box 12.7 considers how physical writing speed affects cognitive processing.

Box 12.7 Practical Application: Taking lecture notes – speed predicts quality

As a student, your writing skills are tested thoroughly throughout your studies, with examinations, essays and practical reports all demanding the type of high-level planning and execution described in the Hayes and Flowers model above. But what of writing tasks such as lecture note taking? How might aspects of that writing task affect learning?

Peverly (2006) summarizes the key skills involved in taking lecture notes. First, information must be held in verbal working memory and this is subject to capacity restrictions (see Chapter 5). Second, the key points must be selected from the information held in working memory, and third, those key points must be transcribed before they are forgotten; this requires efficient writing. All of this must take place while attention is maintained on the ongoing lecture. Although many individual differences come into play here (including working memory capacity, verbal ability, etc.), Peverly (2006) was most interested in writing speed itself. If you can write the key points down quickly, will later test performance be facilitated?

Peverly and colleagues had students listen to a lecture on problem solving and take notes. Measures of transcription speed, working memory, spelling, and identification of salient information were also taken. Their data showed that faster handwriting speed was associated with higher quality lecture notes. By practising the basic processes that allow us to take notes, we increase note-taking speed, and this frees up working memory for attending to the higher level processes involved in selecting salient information, following the argument being made in the lecture, integrating the current information with previous points made, and so on. As Peverly summarizes, 'the best way to enhance the efficiency of a limited-capacity processing system is through instruction and practice, especially of basic skills, so that the capacity of working memory can be devoted to the higher order skills necessary to achieve academic goals' (2006, p. 209).

These data suggest that interventions designed to improve transcription fluency may lead to improved lecture note taking and thereby improved test performance.

Summary

In this chapter we have considered the nature and components of language and the cognitive processes involved in the production of speech. Language is our principal means of communication and seems to be uniquely human. While language shares properties with other animal communication systems, no animal system has all its features. The special features of language are productivity, displacement and duality of patterning.

The components of language are phonemes, morphemes, syntax and discourse. The basic sound units of a language are phonemes; its meaning units are morphemes. Sentences are composed of morphemes and are structured using syntax.

Speech production involves four main stages. Stage 1, conceptualization, prepares a thought for conversion into language. The second stage involves the formulation of a linguistic plan. Formulation also involves syntactic planning; during this sub-stage the order in which the selected words will be output is decided. The third stage involves the articulation of the plan. During this stage the sounds for the word are accessed and articulated. In a final fourth stage, the output is monitored so that corrections can be made if errors occur. Models of speech production differ in terms of the degree of modularity and interaction said to occur between processing levels.

This chapter examined speech errors and their contribution to our understanding of speech production. Speech errors occur in a number of types and are not random. They support the idea that the production of speech involves a number of distinct stages.

This chapter also examined the language deficits that follow brain injury in adults. The patterns of deficit in Broca's aphasia, Wernicke's aphasia and anomia suggest a dissociation between syntactic/output and semantic/comprehension processes in language processing.

Finally, we considered language production in the form of writing, and the three stages of the writing (as composition) process: planning, translating and reviewing.

Review questions

1 What are the main features of human language?
2 What do the acquired disorders of language contribute to our understanding of normal speech production?
3 How does the analysis of speech errors contribute to our understanding of normal speech production?
4 What are the key differences between modular and interactive accounts of speech production?
5 How do the processes of writing differ from those of speech production?

FURTHER READING

Brown, R., & McNeill, D. (1966). The 'tip of the tongue phenomenon'. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*, 73–111.

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, *130*(5), 1432–1441.

Harley, T. A. (2010). *Talking the talk: Language, psychology and science*. Hove: Psychology Press.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*, 88–96.

Preview questions

1. How does language differ from other forms of communication?
2. What do slips of the tongue tell us about the processes involved in speech production?
3. How might brain damage affect speech production?
4. How might we define 'language'?
5. What kinds of processes are involved in writing?

© mrPliskin/iStock

INTRODUCTION

In 1970 in California, a case of child neglect was discovered that was to have a profound impact on our understanding of language development. A girl, named in the literature as 'Genie', had been isolated from the age of 20 months until she came to the attention of social services at the age of 13 years and 7 months. When found, Genie was undersized and severely malnourished. She had painful calluses from being physically restrained over long periods. Throughout her years of isolation, Genie had not been spoken to and she had been beaten when she made a noise. She had not been exposed to language – she had spent many years locked in a room at the back of the family home where she was not able to overhear her family's conversations, and where she was not exposed to language sounds from radio, television or other sources (Curtiss, 1977).

When Genie was found, she did not speak and seemed to understand no more than a few words (Rymer, 1992). Once Genie was taken into care and was exposed to language, in some ways her language development seemed to proceed as it would for a younger child who had typical exposure to language (Curtiss, 1977). She progressed from single words, to two-word and then three-word combinations, and she rapidly acquired vocabulary. However, her language development showed some significant deviations from the normal pattern. Genie had a vocabulary of more than 200 words before she began to combine them, whereas children typically combine words earlier. The word types evident in her early vocabulary also differed from the normal pattern. For example, while most children's early vocabulary consists of basic class words (e.g. 'dog', 'cat'), Genie's vocabulary development showed an emphasis on colours and numbers, shape and size terms, and basic (e.g. 'dog'), superordinate (e.g. 'animal') and subordinate (e.g. 'Labrador') category words (Curtiss, 1981; Fromkin et al., 1974). Genie seemed to seek out words that would allow her to differentiate between similar objects (e.g. pen versus pencil) rather than acquiring labels for a category of object. By contrast, a typically developing child might initially use 'pen' for pens, pencils, crayons and other objects of similar shape, a pattern referred to as over-extension.

Syntax refers to the rules governing the ways words can be combined to create meaningful sentences.

Content words are words that provide meaning to the sentence; these contrast with function words which do the grammatical work of the sentence.

Language production refers to a number of processes by which we convert a thought into language output, in the form of speech, sign language or writing.

Social cognition refers to the ways in which people make sense of themselves and of others in order to function effectively in a social world.

The most striking feature of Genie's language reflects her problems developing **syntax** or grammar (see Table 12.1), evident in the ways she combined words. Curtiss (1981) described Genie's sentences as '. . . the stringing together of **content words**, often with rich and clear meaning but with little grammatical structure' (p. 21). Some examples of Genie's utterances illustrate this feature of her language: 'I like hear music ice cream truck'; 'Think about Mama love Genie'; 'Dentist say drink water'; 'Applesauce buy store' (Curtiss, 1977). Genie's case reflects three key issues in language acquisition. First, her failure to fully acquire language suggests that there may be a critical or sensitive period for language acquisition, and particularly for grammar development; if the child is not exposed to language in a social context within this period, normal development is constrained (see Lenneberg, 1967). Second, her language reflects the dissociation between the acquisition of vocabulary and the flexible use of this vocabulary to form novel sentences. Third, her case suggests that language acquisition, like many cognitive functions, relies on interplay between input from the environment (nurture) and biological makeup (nature).

This chapter examines the nature of language and the cognitive processes involved in **language production**, with a focus on speech. Language is a quintessentially human ability and the ability to communicate our thoughts to others through language is fundamental for **social cognition**. Language also shapes mental representation and thinking (e.g. Crystal, 1997). Once acquired, language is 'fundamental to all distinctly human thought and consciousness' (Donald, 1999, p. 139). It is therefore important for cognitive psychology to study the processes involved in speech production and to try to understand how a thought is turned into spoken words.

Two basic stages in speech production have long been recognized: the formulation of a thought and its conversion into speech (e.g. James, 1890). However, cognitive psychology

Table 12.1 Levels of linguistic analysis

Level	Refers to:
Semantics	The level of meaning in language
Syntax	The rules by which words are combined to make meaningful sentences
Morphology	The rules by which words are constructed and modified
Phonology	The sound units within a language

remains relatively poorly informed as to the precise nature of the processes underlying language production. As we will see in the next chapter, greater progress has been made towards understanding language comprehension, while less research has addressed the processes underlying production (MacNeilage, 1999). Methodological constraints provide one reason for this bias: it is difficult to control experimental stimuli in order to study language production. When examining language comprehension, we can manipulate the words, sentences or other stimuli that are presented to research participants and measure the effect on comprehension. Comprehension follows on the presentation of the stimuli. But production proceeds from cognition to motor output, and it is a far more difficult task to control or inspect the content of someone's thoughts. Speech production proceeds in a top-down manner, that is, it is **conceptually driven** (see Chapter 2).

Conceptually driven or top-down processes reflect the influence of higher order cognitive processes such as thoughts, beliefs and expectations.

In spite of the methodological challenges posed by the topic, our understanding of speech production has seen substantial development in recent years. Knowledge of speech production derives from a number of sources, involving, for example, experimental methods, computational modelling, neuroscientific methods and neuropsychological case studies. In the present chapter we focus on learning about language production by examining what happens when the system fails. Two types of system failure are considered. First, we consider speech errors; slips of the tongue and other speech errors reveal much about the processes underlying speech. Errors can be induced experimentally or recorded from spontaneous speech. Second, the effects of physical damage to the areas of the brain responsible for language production will be examined.

Let's start by considering what language is: How might we define 'language'?

LANGUAGE AND COMMUNICATION

Language is our principal means of communication and forms the basis of the majority of social interactions. **Communication** can be fairly readily defined as any means by which information is shared (e.g. Field, 2003) or as a process whereby 'a source encodes and transmits a signal, which is detected by a receiver and decoded into meaningful terms' (MacFarland, 1999, p. 387). Many definitions of language would include its use in communication as a core feature, but clearly language goes beyond communication. The information sharing function of language may be a relatively minor role; Aitchison (1996, p. 25) suggests that language has been particularly important for human evolution because it promotes social bonds and social interaction and because it provides an effective means of persuading others.

Communication refers to any means by which information is shared.

There are two ways in which we can use language to communicate. One way is through writing. Written language is a new (and arguably humankind's greatest) invention; the earliest evidence of writing dates to about 5000 years ago. Writing developed from a number of distinct systems originating in different parts of the world and this is reflected in the considerable cross-language variation in scripts today (see Chapter 13). Writing involves converting thoughts or speech to print. In contemporary society, reading and writing are essential skills and anyone who fails to acquire them, for whatever reason, is at a considerable disadvantage. Writing also plays a vital role in language survival, by allowing a record of the language to be retained across generations: a language without writing is unlikely to survive.

The principal way we use language is through speech (or a related mode of output such as sign language). This aspect of language has been a feature of human cognition for tens of thousands of years, and is without parallel in the animal kingdom. Spoken language is found in all human groups and would seem to be qualitatively different from the communication systems of other species.

People also communicate non-verbally. Non-language vocalizations (e.g. grunts, groans) can convey information, and gestures can supplement or substitute for spoken language (see Jacobs & Garnham, 2007; McNeil, 1992). There are many speculative accounts of the origins of human speech. Some accounts highlight the interaction between spoken language and gesture use (e.g. Corballis, 2003). Gesture is so closely tied to human language that we continue to gesture even when we cannot be seen; it is common for people to gesture as they communicate over the telephone, for example (Bavelas et al., 2008; see Chapter 8 for more on gesture and embodied cognition). Subtler non-verbal signals such as body language and tone of voice (see Chapter 2) also communicate to others, whether we are aware of this or not.

Languages vary on a number of dimensions, but also have features in common, an issue we explore next.

LANGUAGE UNIVERSALS

Estimates regarding the number of languages in use worldwide vary considerably, depending on the criteria used to count speech systems as distinct languages. There are about 6000 languages (e.g. Comrie 1989; Crystal, 1998; Krauss 1992; Moseley, 2007) in use worldwide, and this figure is decreasing. Many of these languages are close to extinction, with a very small number of speakers and few or no child speakers. For example, Krauss (1992) noted 187 indigenous languages in North America, only 20 per cent of which were still being learned by children. In fact, just 4 per cent of the world's languages is spoken by 96 per cent of the world's population, placing many languages on the 'endangered' list (Crystal, 2000). Krauss (1992, p. 7) estimated that as few as 10 per cent of the world's languages will remain in a hundred years, with minority languages facing increasing pressure from the dominant, majority languages (see also Crystal, 2000).

Languages vary in the number and type of sounds used, in basic word order, in the size of their vocabularies (reflecting the number of items in the **lexicon**) and in their rules for sentence construction. However, all are capable of expressing complex and new ideas: there are no 'primitive' languages. Though the precise way in which concepts are expressed may differ across languages (e.g. Boroditsky, 2001; see Chapter 13), the expression of complex ideas is evident in all languages and in all human groups.

Mental lexicon is our store of knowledge about words and their uses.
Linguistic universals are linguistic features said to be found in all languages.

Languages have some key features in common, though it proves to be a difficult task to identify a set of **linguistic universals**, that is, features that are shared by all languages. Aitchison (1996, p. 177) lists the following features as 'absolute universals' and acknowledges that even these are problematic. According to Aitchison, all languages:

- have consonants and vowels;
- combine basic sounds into larger units;
- have nouns (words for people, places and objects, e.g. book);
- have verbs (words that represent actions or 'doing', e.g. to read);
- can combine words in meaningful ways;
- can express who did what to whom;
- can express sentences as negatives;
- can express sentences as questions;
- are structure-dependent, that is, involve a syntactic structure or grammar;
- allow recursion (the use of a rule within itself, allowing, for example, embedded sentences).

There are, as Aitchison points out, immediate problems with this listing. Sign languages are language, though they do not use a system of vowels and consonants. Some languages do not reliably distinguish between classes of nouns and verbs. Some nouns can represent actions (e.g. 'destruction') and nouns can be used as verbs ('to text'). Table 12.2 provides a summary of the parts of speech, such as nouns and verbs. While all spoken languages are based on combinations of vowel and consonants sounds, the precise set of sounds varies considerably across languages (see the section on phonology to follow). MacNeilage (1999) identifies the syllable, that is a vowel and consonant combination, as a universal unit of speech. Some languages use other sounds in addition to vowels and consonants. For example, in **tonal languages**,

Tonal languages use changes in tone to alter the meaning of the word.

Table 12.2 The lexical categories (word classes or 'parts of speech') of English. See *The Cambridge Grammar of the English Language* (Huddleston and Pullum, 2002).

Category	Description
Verb	A very large class of words among which are many denoting actions (*run*, *jump*, *swim*) or states or experiences (*feel*, *see*, *think*). Auxiliary verbs mark tense, aspect and modality (*has eaten*, *is eating*, *may eat*). Verbs that take a direct object (as in *The dog ate the bone*) are called transitive verbs. Others are intransitive (*The dog slept*).
Noun	An extremely large class of words including the words denoting people, places, things, or abstract ideas in the broadest sense. The nouns include not only *desk*, *book*, *football*, *house*, etc., but also *absence*, *thought*, *possibility*, *failure*, etc. Proper nouns begin with a capital letter and name particular people, places, or things: *Barack Obama*, *London*, *Disneyland*. The pronouns are a special subclass of the nouns, used for referring to things without naming them (often because they are named elsewhere in the context): *he*, *she*, *I*, *you*, *they*, *it*.
Determinative	The fairly small class of determinatives includes words usually used with nouns to specify definiteness or quantity: *the*, *a*, *some*, *all*, *every*, *many*, *much*, *several*, *this*, etc. Sometimes they occur without nouns (*Some may disagree*; *I like this*), or in other uses (*So much the better*).
Adjective	The adjectives are a large class of words that typically modify or qualify the meaning of nouns, as *brown* does in *the brown dog* or *Our dog is brown*: words like *happy*, *cloudy*, *intelligent*, *mysterious*, *cool*, and many others.
Adverb	The great majority of the adverbs are formed from adjectives by adding *ly*: *nice* is an adjective and *nicely* is an adverb. They very commonly modify words other than nouns, providing information about where, when, how, or to what extent things happen: *locally*, *recently*, *awkwardly*, *extremely*, Adverbs unrelated to adjectives include *soon*, *quite*, *too*, *always*, *seldom*.
Preposition	Prepositions often precede noun phrases to indicate spatial or temporal relationships, as in *the book is on the shelf*. The prepositions include *after*, *at*, *before*, *by*, *despite*, *during*, *except*, *for*, *in*, *into*, *of*, *on*, *over*, *since*, *through*, *to*, *under*, *with*, *without*. some occur before clauses (*after we met*), and some occur without anything following (*Go right in*).
Subordinator	There is a very small class of words called subordinators that function to mark clauses embedded in other clauses: *that* as in *that nobody loves me*; *whether* as in *whether anybody loves me*; *for* as in *for someone to love me*.
Coordinator	There is a very small class of words called coordinators, used to link clauses or phrases together: *and*, *or*, *but*, *nor*, and a very few others. (The subordinators and coordinators are traditionally classed together and called 'conjunctions'.)
Interjection	An interjection is a word that interrupts a sentence with an immediate expression of emotion or sentiment: *yikes*, *oops*, *shush*, *ouch*, etc.

altering the tone of expression communicates meaning. In English, and non-tonal languages generally, we can change the tone of the utterance for emphasis, or to convey emotion, but doing so does not alter the meaning of the word. In a tonal language, the tone carries meaning. In Mandarin, for example, a language with a relatively small number of syllables, 'ma' can mean 'mother', 'horse' or 'scold', if the tone is even, falls then rises, or falls, respectively (Ladefoged, 1993, p. 255).

Some languages use unusual classes of consonant sounds. There are about 30 languages in Southern Africa (e.g. the Xhosa, Khoikhoi and Sosotho languages) which use a 'click' sound (like the 'tut tut'/ 'tsk' sound in English). These clicks may originally have aided communication while hunting, as they mimic natural environmental sounds and therefore would not have startled prey (Knight et al., 2003). Sounds that qualify as meaningful differ substantially across languages. Even in languages with many features in common (e.g. German and English) the precise set of speech sounds varies. Given this diversity, an approach to defining language based on broad design features may prove more fruitful.

Hockett's design features for human language

Scan to watch a video about animals learning language

Charles Hockett's (1960) set of 16 design features for human language was formulated with the aim of identifying properties that are unique to human languages and differentiate them from other animal communication systems. While animal communication systems share some of the following features, only human language demonstrates the full set. The design features treat speech as the standard mode of expression.

1 *Vocal–auditory communication channel*: Languages normally transmit information via spoken sound with the sender speaking and the receiver hearing the spoken signal (sign systems use the tactile-visual medium to similar effect). Many animal systems use vocalizations as the means of communicating while other animals use non-vocal means to communicate. For example, a honeybee uses a 'figure of eight' shaped dance to signal information about the location, distance and quality of a food source (Von Frisch, 1962).

2 *Broadcast transmission and directional reception*: The speech signal is transmitted out from the source (the speaker's mouth) and is localized in space by the receiver.

3 *Rapid fading*: The spoken message fades after production, unlike, for example, written language, which can be inspected over time.

4 *Interchangeability*: The sender can also be a receiver and vice versa. The speaker role is interchangeable and not fixed.

5 *Feedback*: The speaker has access to the message and can monitor its content. This allows us to monitor and correct errors or slips in spoken language and, as Hockett pointed out (1960, p. 6), this kind of cognitive access forms the basis for the internalization of language as verbal thought.

6 *Specialization*: The energy expended in producing the message does not alter the meaning of the message. Whether we whisper or shout the utterance, the meaning of the words remains the same, although we can change the emphasis by altering the vocal energy, or indeed communicate non-literal meaning using a change in tone (e.g. sarcasm).

7 *Semanticity*: Sounds within speech refer to objects and entities in the world: words have meaning. There is an element of semanticity in some animal calls. For example, vervet monkeys use a system of predator alarm calls with distinct calls for snakes,

eagles and leopards (Seyfarth et al., 1980). The calls are more likely to be made in the presence of other monkeys and, in particular, in the presence of kin (Cheney & Seyfarth, 2005). This use of specific calls by animals is referred to as **functional reference** and the information contained in the signal allows the signaller's conspecifics (members of same species) to react appropriately. However, there is little flexibility in their use.

Functional reference refers to the use by animals of a specific call to stand for a specific object or threat.

8 *Arbitrariness*: The relationship between the spoken word and its referent in the world is arbitrary. Apart from a small number of onomatopoeic words (e.g. *hiss, buzz, cuckoo* and *murmur*), the form of the signal (the way it sounds) does not relate to its meaning directly.

9 *Discreteness*: The speech signal is composed of discrete units. Vocabulary is built up from smaller meaning units and the meaning units in turn are built from the basic sounds of the language. These sounds are perceived categorically. Change the /p/ sound in 'pin' to a /b/ sound and you have a different word with its own meaning, 'bin' (we follow the convention here of representing sounds by a letter between forward slashes, e.g. /p/).

10 *Displacement*: We can use language to refer to things that are displaced from the present situation, either in time or space. By contrast, animal systems tend to be tied to the current context. In the vervet monkeys' signalling system mentioned above, the issuing of an alarm call is triggered by the presence of a perceived threat but the calls cannot be used flexibly outside of that context. There is no way to communicate 'I saw that eagle earlier' or 'Bob says he saw a leopard' (see Harley, 2010).

11 *Productivity*: Language allows us to create novel utterances. This aspect of language is a fundamental distinguishing feature of human language and its basis lies in syntax (Chomsky, 1986). Productivity is also referred to as openness or generativity. From the earliest stages of language development, speech is characterized by novelty. We do not just repeat back speech we have heard; we say things in new ways. Likewise, we can understand novel sentences that we will not have encountered before.

12 *Cultural transmission*: A language is learned through interaction with more experienced users of the language within a verbal community.

13 *Duality (of patterning)*: Meaningful elements are created by combining a small set of meaningless units. For example, the 40 or so sounds of the English language are as and of themselves meaningless; however, they combine in meaningful ways to allow us to utter all the words in the English language.

14 *Prevarication*: Language can be used to deceive and lie. Furthermore, the messages we create may lack an obvious meaning.

15 *Reflexiveness*: We can use language to communicate about language. The sentences you are currently reading are an example of this property of language. Animal systems lack this feature; as Harley (2008) comments, 'bees will never dance a book about the psychology of the bee dance' (p. 57).

16 *Learnability*: A language can be learned by a speaker of another language.

These features are not independent, as Hockett noted. For example, semanticity and arbitrariness are related: words have meaning, they refer to something in the world (semanticity) and the relationship between the sound of the word and the thing it refers to is not physically direct (arbitrariness).

The design features apply to spoken language and do not apply fully to sign languages (which it is generally agreed show similar linguistic properties to spoken language) or to written language. However, they provide a useful general way of differentiating animal communication systems from language. While animal systems have some of Hockett's properties for language, the full set of design features is only found in language.

COMPONENTS OF LANGUAGE

Language is a structured system which uses a finite set of sounds to construct words, sentences and ultimately conversations. The components of language, from the smallest to the largest parts, are phonemes, morphemes, syntax and discourse. In order to understand language, it is useful to consider these components independently; we start with the sounds of language, phonemes.

Phonemes

Phonemes are the basic sounds that make up speech within a language and the term phonology refers to the system of sounds in a language. There are about 100 basic sound units or **phones** (as listed in the International Phonetic Alphabet) that can be used to make up words. This represents the full set of available sounds; the study of these raw sounds is called **phonetics**. No one language uses all these sounds; instead, languages use a subset of phones, and languages vary considerably in the number of sounds used. The sounds within a language are called **phonemes**. These are the smallest meaningful sound units in a language. The number of phonemes within a language varies – there are 44 in (British) English, while some Polynesian languages (e.g. Hawaiian) have as few as 12, and there are over 140 in the African language Khoisan (Chierchia, 2001). Table 12.3 shows the number of phonemes in a selection of European languages.

Phones are the basic speech sounds.
Phonetics is the study of speech sounds.
Phoneme is the smallest meaningful sound unit within a language.
Allophones are phonetic variants of the same phoneme.

Some basic vowel sounds occur in all languages, but consonants can be used differently and are perceived differently. For example, in English, there is no perceived difference between the aspirated (i.e. said with a puff of air) /p/ sound in *pin* and the unaspirated /p/ in *spin*, but in Thai these are two distinct phonemes. Similarly, in English the /t/ sounds in *tea* and in *trip* are different phones, but these are treated as one phoneme; if you substitute the *t* sounds, the pronunciation may sound a little odd, but it is still the same word. Phonemes also change as a function of the surrounding sounds in words and in sentences, an effect referred to as co-articulation. Different phones that are treated as the same phoneme within a language are called **allophones**. Phonemes therefore do not correspond directly to physical sounds; rather they are 'abstract representations of the phonological units of a language, the units used to represent the forms of word in our mental lexicons' (Fromkin et al., 2003, p. 285). A phoneme is a rather subjective category that is recognized as meaningful by the speakers of a language, but is not necessarily constant as objectively measured (see Chapter 13).

Table 12.3 The number of phonemes in a selection of European languages (from Tambovtsev & Martindale, 2007).

Language	Number of phonemes
Finnish	56
German	51
Italian	49
English (British)	44
Swedish	41
Dutch	39
French	36
Norwegian	36
Greek	25
Portuguese	25
Spanish	24

Source: SKASE Journal of Theoretical Linguistics

The tendency to perceive the difference between two allophones decreases with age (e.g. Iverson et al., 2003), suggesting a critical period beyond which the adult is 'tuned' to the sounds of his or her native language. While a child can discriminate between the full set of phones, an adult generally cannot; a child who is exposed to a second language can acquire native-like ability and accent, while the parents, coming late to the language, may struggle to acquire it and never acquire a native-like accent (see Bongaerts et al. 1995; Dewaele, 2009; Singleton, 2001). For example, the difference between /l/ and /r/ sounds in English is not readily discriminated by Japanese adults, for whom /l/ and /r/ are treated as allophones (Ingram, 2007). While young children can appreciate the difference, by adulthood this ability is reduced. This reduction in discrimination with linguistic experience may serve to reduce the ambiguity in the incoming speech signal, facilitating language comprehension, an issue we return to in Chapter 13.

Phonological and **phonotactic rules** describe which sounds can go together in a given language. For example, in English, a /t/ sound does not follow a /b/ sound and [ng] can occur at the end of the word (as in 'king') but not at the beginning. These rules differ across languages so that sounds that are 'natural' and easy to produce and discriminate in one language may not be so easy for adult speakers of another language. Speech segmentation relies on knowledge of word boundaries using information about phonotactic probabilities in a language (see Chapter 13).

Phonotactic rules stipulate which combinations of sounds are 'permitted' in a language.

Changing a phoneme within a word will change the meaning of that word, for example *bat* and *pat*. This is an example of a minimal pair, so called because the words differ by just one phoneme. All spoken languages use vowels and consonants, but, as outlined above, the exact set used varies across languages. Maddieson (1984) examined a 300 language sample and found that the number of consonants varied from 6–95 sounds, with a mean of 23, while vowel sounds varied from 3–47, with a mean of 9.

Morphemes

Morphemes are the meaning units of a language. They are the building blocks of words. A single word may consist of several morphemes. The term **morphology** refers to the study of the rules in a language according to which words can be constructed. Morphology can be considered as a special case of syntax (see below).

In English, regular plural nouns are created by adding –s to the end of a word, for example, *one car* but *two cars*; *one horse*, *two horses*, etc. In these examples, there is one morpheme in the singular forms (*car*) but two in the plural (*cars*), the stem or root word (car) and the plural suffix inflection (-s). The noun *car* is an example of a **free morpheme**, as it can occur on its own, whereas the plural form –s is a **bound morpheme**, because it does not carry meaning unless it is attached to a free morpheme. Here, the –s is an example of an inflectional morpheme; it serves a grammatical function but does not change the syntactic category of the word to which it is attached (*car* is still a noun when the –s is added to make *cars*). Similarly, the verb endings –ed and –ing are inflectional morphemes. Some bound morphemes, like –ify, –ish, –able and –ment, are derivational morphemes, as they create new words with new meaning when added to a stem. They can change the grammatical category of the word. For example, the verb *develop* becomes a noun when you add –ment to give *development*. Words can be altered by adding a morpheme to the start of the word (a prefix) or to the end (a suffix) and language-specific rules govern the ways in which words can be altered. The verb *depend* becomes *dependence* when we add the suffix –ence, meaning 'condition' or 'state'; adding the prefix –in, meaning 'not', yields *independence*, and so on.

Morphemes are the meaning units of language.
Morphology is the level of linguistic analysis concerned with morphemes and their role within words.
Free morpheme is a morpheme that can stand alone as a word.
Bound morpheme is a morpheme that cannot form a word on its own, but forms a word when attached to a free morpheme.

Such alterations apply only to content words (e.g. nouns, adjectives and verbs); these are open class words which can be altered or invented as usage changes. In some languages inflections on content words can be particularly informative. In Hungarian, for example, the morpheme at the end of the word indicates the word's role in the sentence, and codes whether it is a direct or indirect object. For example, consider these sentences in English and in Hungarian (example from Hoff, 2005):

The boy gave a book to the girl.

A fiú egy könyvet adott a lánynak. (The boy a book gave the girl.)

Here, 'book' is the direct object in the sentence (it is given), while girl is the indirect object (the book is given to the girl). While in English we tend to rely on word order, in most cases, to understand the role of the word in the sentence, in the Hungarian sentence above the morphemes *et* at the end of 'könyvet' and *nak* at the end of 'lánynak' give the role of the word. In this example, content words are altered to indicate the word's role in the sentence; in grammatical terms the inflections are accusative and dative case markers, respectively.

Function words provide grammatical structure that shows how content words relate to each other within a sentence.

Function words, the words that do the grammatical work of a sentence do not change (prepositions, for example; see Table 12.2) – they are a closed class of morphemes. As we will see, content words and function words are to some extent treated differently in language processing.

Semantics and the lexicon

Morphemes make up words, which in turn make up our vocabulary. Our knowledge of words and their meanings are stored in a kind of mental dictionary called the mental lexicon. The lexicon is a part of the semantic memory system (see Chapter 5). It holds our store of words and associated knowledge and links words with our general knowledge about concepts and the world. As adults, we have a store of tens of thousands of words, from which we normally have immediate access to target words as we construct a sentence. Only occasionally will we experience difficulty in calling a target word to mind, a temporary failure referred to as the tip-of-the-tongue effect (more on this later in this chapter).

It is difficult to estimate the size of an adult's vocabulary; some studies suggest that adults know about 70,000 words (e.g. Bock & Garnsey, 1998; Nagy & Anderson, 1984). Miller (1977) estimated that the average English speaking college student had a vocabulary of about 150,000 words; subsequent estimates put the figure closer to 20,000 word families, where a word family consists of a base word, and inflected and derived forms (e.g. Goulden, Nation & Read, 1990). The range of estimates reflects the difficulty in accurately measuring vocabulary and the importance of defining the unit of measurement.

Words are symbols; they are meaningful sounds and generally have a particular referent. A word might be defined as 'the smallest unit of grammar that can stand on its own as a complete utterance' (Crystal 1997, p. 440); in writing, words are generally separated by spaces. People also use other meaningful sounds, though not all are words. For example, we might use a groan to signal disagreement or a yawn to signal boredom, but these are not words. A few words are not referential, that is they have no clear referent – greetings and social conventions (e.g. saying *hello*) for example.

Word is the smallest unit of grammar that can be meaningfully produced on its own; it can consist of one or more morphemes.
Semantics is the study of meaning.

The question of what words mean and how they relate to each other raises some complex issues. **Semantics** refers to the meaning of words and morphemes and the relationship between the words we use and the objects they refer to in the world.

Syntax

We construct novel sentences when we speak; we do not generally repeat back or 'parrot' previous productions. Imagine you are telling a story to a friend when another friend joins the group and you have to start your story over. The chances are, though the meaning or gist of your story will not change, the exact sentences you use will differ. This reflects the **productivity** of human language; we do not rely on rote or stock phrases, or on memory for practised utterances. Instead we create new sentences as and when we need them. This is evident from the earliest stages of syntactic development in young children.

Productivity of language refers to the ability to generate novel utterances.
Phrase is a group of words referring to a particular idea.
Slang describes an informal pattern of speech that is considered to be 'non-standard'.

Two aspects of the language system allow us to use language productively: syntax and morphology. The term syntax describes the rules that determine the construction of **phrases** and sentences in a language. It relates to grammar but the term syntax is used more often than grammar, to differentiate it from the notion of 'prescriptive' grammar. Prescriptive grammar reflects conventions for sentence construction and is based on tradition and language prestige rather than actual language use. For example, split infinitives ('*to boldly go* where no-one has gone before') and prepositions at the end of sentences ('prepositions are not good words to end a sentence *with*') violate conventions of prescriptive grammar, but are often found in everyday speech.

Similarly, **slang** may not always meet with approval, but could nevertheless be syntactically correct: the ambiguous 'I don't know nothing' might not be considered 'good form', but the same sentence would never be uttered as 'know I nothing don't'. The study of syntax reflects descriptive grammar, that is, it reflects how language is used.

Sentences follow a hierarchical structure and are made up of two parts: a noun phrase (NP) which contains a noun, often the **subject** of the sentence, and a verb phrase (VP) which contains the verb and conveys the 'action' of the sentence. For example, in the sentence 'Sarah drank the coffee', Sarah (the subject of the sentence) is the NP and 'drank the coffee' is the VP. English, French, German and related languages use a subject-verb-object or SVO word order, that is, in a declarative (active voice) sentence the subject (or agent of the sentence) comes first, followed by the verb and then the **object** of the sentence. The order in which the words occur determines the meaning of the sentence; to use Pinker's (1994) example, 'dog bites man' is not newsworthy, but 'man bites dog' is (p. 83). The most common word orders are SOV and SVO (Greenberg, 1963) and although there are examples of the six possible types (SOV, SVO, VSO, VOS, OVS, OSV), OVS and OSV are extremely rare (Dryer, 2005; Song, 2001). Some languages have more flexible words orders; for example, Japanese is mainly SOV and Russian is SVO but both languages can use other word orders because of their use of case markings. This agent-first bias in world languages is not restricted to spoken languages. It is also found in the 'homesign' produced by deaf signers with hearing parents (Goldin et al., 1990). It also appears in second languages acquired without explicit instruction (Klein & Perdue, 1997).

Subject of a sentence is the word or words that gives what the sentence is about or performs the action.
Object of a sentence is the word or words that receives the action, or is acted on, by the subject of the sentence.

One key property of syntax underlies the productivity of sentence construction. **Recursion** refers to the repeated application of a rule and, using recursion, the same rule can be applied again and again to create a novel

Recursion refers to the ability to extend sentences infinitely by embedding phrases within sentences.

utterance. Recursion has been argued to be an essential property of human language (e.g. Chomsky, 1986). Embedded sentences make use of this property, and sentences can in principle (though not generally in practice) be extended indefinitely. For example, the English language nursery rhyme 'The house that Jack built' is an example of a cumulative rhyme using recursion:

> This is the house that Jack built.
>
> This is the malt that lay in the house that Jack built.
>
> This is the rat that ate the malt that lay in the house that Jack built.
>
> This is the cat that killed the rat that ate the malt that lay in the house that Jack built.
>
> This is the dog that worried the cat that killed the rat that ate the malt that lay in the house that Jack built . . .

Recursion would seem to be a resilient property of human language as even young children who have been deprived of language input retain the ability to use recursion (Goldin-Meadow, 1982). The extent to which recursion is uniquely human has been challenged, however. For example, songbirds have been shown to be sensitive to recursion and can classify novel patterns accordingly and reliably reject 'ungrammatical' patterns (Gentner et al., 2006).

Discourse

Discourse refers to multi-sentence speech and includes dialogue, conversation and narrative. **Pragmatics** refers to the understanding of the communicative functions of language and the conventions that govern language use.

Discourse refers to multi-sentence speech and includes dialogue, conversation and narrative. At this 'higher' level of language function, the social conventions that affect language processing become increasingly relevant and people rely on schemas (see Chapters 5 and 7) in order to process language. **Pragmatics** refers to the understanding of the communicative functions of language and the conventions that govern language use. At the level of discourse, the function of language in communicating directly and indirectly comes to the fore. A distinction is made between linguistic competence, which refers to our ability to construct sentences, and communicative competence, which refers to our ability to communicate a message effectively (Hymes, 1972). Language can be perfectly well formed, but if we fail to appreciate the social conventions governing its use, we may not communicate as we intended.

Effective discourse is based on a shared understanding between those engaging in a conversation. For example, if two people are conversing and one asks the other a question, there is an implicit agreement that the response will be related to the question. Similarly, participants in a conversation are expected to adhere to the topic of the conversation. If someone wishes to deviate from the topic or to change the subject, it is customary to signal this change of focus, by prefacing the utterance with 'by the way . . .', for example.

Conversations require turn-taking and cooperation and participants follow a set of implicit social conventions. A variety of verbal and non-verbal signals serve to regulate the conversation by indicating who speaks when and for how long. These turn-taking cues act to minimize overlap between speakers and reduce gaps or silences in conversation. Conversational turn-taking has several features (Sachs et al., 1974). One party speaks at a time; the person speaking changes. The duration of a turn is not predefined; the order of turns also varies. Transitions between turns are coordinated; overlap is minimized. These patterns hold in the absence of face-to-face information, for example, in telephone conversations (De Ruiter et al., 2006). Despite the differences in linguistic features across languages, there are striking universals in turn-taking patterns, as is explored in Box 12.1.

Box 12.1 Research Close Up: Cross language universals in conversational turn-taking

Source: Stivers, T. et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 10587–10592.

INTRODUCTION

Stivers et al. set out to examine the extent to which there are cultural differences in turn-taking in everyday conversations. They tested two opposing hypotheses. The 'universal system' hypothesis predicts little cross-linguistic variability and predicts that most languages will use a 'minimal-gap minimal-overlap' convention, as in found in English. On the other hand, the 'cultural variability' hypothesis, based on anthropological accounts, holds that turn-taking practices differ considerably across languages and cultures.

METHOD

Stivers et al. (2009) analysed video recordings of informal conversations in 10 languages from five continents. The languages varied in structural properties (e.g. word order, grammar) and were drawn from different cultures, ranging from hunter–gatherer groups to large-scale industrialized societies. All conversations were spontaneous, informal conversations, with 2–6 participants. Questions and responses were timed, coded for their form and function, and coders judged whether the responses were delayed.

RESULTS

Striking similarities emerged across the languages supporting a 'minimal-gap minimal-overlap' norm. While there was some variation across languages, the mean response time for a turn transition was very similar across languages (see Figure 12.1). Of the languages examined, Danish had the slowest response time on average (+469 milliseconds), while Japanese had the fastest (+7 milliseconds). Italian (309 milliseconds), English (236 milliseconds) and Dutch (109 milliseconds) fell between these extremes. The mean response time across languages was +208 milliseconds, and each language's mean was within 250 milliseconds of the cross-language mean: as Stivers et al. note, this is about the time it takes to say a single syllable in English. In other words, responses tended neither to overlap nor be delayed by more than half a second.

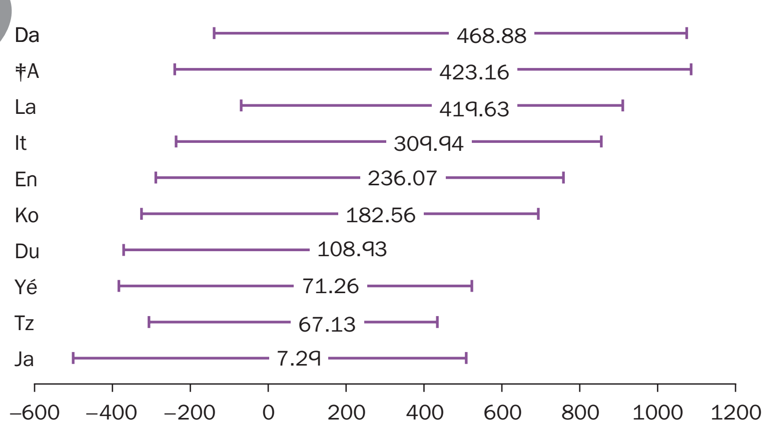

Figure 12.1 Mean time (in milliseconds, shown on the x axis) of turn transitions for 10 languages. Speakers of all languages have an average offset time within 500 milliseconds, while averages vary considerable across the languages. Languages shown along the y axis are: Da, Danish; ǂĀ, ǂĀkhoe Haillom; La, Lao; It, Italian; En, English; Ko, Korean; Du, Dutch; Yé, Yélî-Dnye; Tz, Tzeltal; Ja, Japanese.

Source: Stivers, T. et al. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS*, *106*, 10587–10592.

Furthermore, the factors that predicted the speed of a response were identical across the languages. When visible responses were made (e.g. a headshake, a nod), they occurred faster than speech in all the languages. Confirmation responses were faster than disconfirmation responses. Questions accompanied by a gaze received faster responses than those without, and answer responses (e.g. 'yes') were significantly faster than non-answer responses (e.g., 'I don't recall').

DISCUSSION

These data support strong universals in turn-taking patterns across languages, and suggest a common pattern whereby the gaps between turns, and overlaps, are minimized. While there were some differences between languages, in real terms these differences are very small. The 'long silences' associated with Nordic languages, for example, evident in the Danish response times, amount to a delay equivalent to about one syllable. The data were limited to question–response sequences, however; it would be interesting to see if similar findings apply to other types of interactions.

Grice (1957, 1975) identified a set of four conversation rules or maxims that describe effective conversations and reflect the expectations of listeners. Grice's four maxims are:

1 *Maxim of quantity*: The speaker should provide enough information in order to be understood but not too much information.
2 *Maxim of quality*: The speaker should provide accurate information.
3 *Maxim of relevance*: The speaker should provide information that is relevant to the current topic of conversation.
4 *Maxim of manner*: Ambiguity and vagueness should be avoided.

If the maxims are violated, more cognitive processing is required to determine the response, or the participants may have to backtrack or repair the conversation. Of course, violation of these rules can also form the basis of humour – hyperbole, sarcasm and irony violate the maxim of quality.

Thus far, when we consider the meaning of words we have treated this as what the word *denotes*, that is its literal meaning; a word denotes its referent. We are also able to communicate indirectly, however. If you were sitting next to an open window and someone said to you 'It is cold', the utterance might be taken as meaning 'Can you close the window?' We must also consider the connotations that words evoke. Connotation refers to the non-literal aspects of word meaning and reflects social and cultural factors that affect the literal processing of word meaning. Words can be perceived as having a positive or negative connotation (see Jay & Danks, 1977). If you hear someone say, 'Bob eats like a pig', it is of course literally true, in that the manner of eating (the movements of the jaw, and so on) is similar in many animals; however, it is likely that something else is meant. It is likely that Bob is being insulted and the connotation created by the phrase could mean that Bob is a messy eater or that he is greedy. Similarly, words can attract a positive value through use and this affects understanding of their literal meaning. The word 'natural' is much used in advertising, for example, and its use relies on the fact that people perceive 'natural' as having a value: 'natural' is equated with 'good'. The many entirely natural but life threatening infections that assail humanity do not seem to be evoked by the term, nor do 'natural disasters' spring to mind.

Having considered the constituents of language we now move on to examine how language is put together so that phonemes, morphemes and syntax become discourse. Much of what we know about this process is derived from the study of speech errors.

SPEECH ERRORS

Speech is produced at a rate of about 15 speech sounds and 2–4 words per second (Levelt, 1989) and for the most part is fluent and well formed. Slips and errors in spontaneous speech are quite rare, with some studies finding errors less than once per 500 uttered sentences (Dell et al., 1997; Garnham et al., 1981; see also Levelt, 2001). Error rates are low despite the rapid rate at which sounds must be selected from a production vocabulary of about 20,000 words (Groome et al., 1999). Hesitations and pauses in spontaneous speech are common, however.

Many of the theories of speech production outlined later in this chapter originate from analyses of speech errors. Data about errors come from several sources. The first source involves examining temporary breakdowns in the system's functioning, which occur from time to time, under normal conditions. Speech errors such as slips of the tongue and tip-of-the-tongue effects (when we cannot fully access a target word from the lexicon) are examples of this type of 'malfunction'. Second, errors can be induced in the laboratory, by having people articulate very quickly, for example. Third, the study of acquired brain injury has shown how damage to certain brain areas affects speech and language processing. **Aphasia** is the term used for acquired disorders of language, a topic we return to shortly. First, we look at hesitation and pauses in normal speech.

Aphasia is the term given to a group of speech disorders that occur following brain injury.

HESITATIONS AND PAUSES

Disfluencies such as pauses are more common than actual errors and vary with the situation and the individual. They are a natural characteristic of fluent speech. Estimates suggest that about six in 100 words are affected by disfluency (Fox Tree, 1995). These pauses can be silent or filled (common fillers include *um*, *ah* or *er*; see Maclay & Osgood, 1959, for an early account of their use). Filled pauses occur with less frequency than silent pauses (O'Connell & Kowal, 2004) and may serve to announce a delay in speech (Clark & Fox Tree, 2002). During pauses, the speaker plans the articulation of their next words (Butterworth, 1980). Goldman-Eisler (1968) found that when participants were instructed to speak about a given topic, pause duration accounted for as much as half of the total time (although the method used to arrive at this estimate has been questioned; O'Connell & Kowal, 2004). The use of pauses varies with context, task demands and from individual to individual. Schacter et al. (1991) found that pauses in academic lectures varied with academic discipline, with science lecturers producing fewer pauses than humanities lecturers. This may reflect the more precise terminology employed by scientists; if there are fewer words to choose from, word choice is facilitated. Pauses within **clauses** and sentences would seem to reflect formulation of ideas and word selection (Velmans, 2009).

Disfluency is a hesitation or disruption to the normal fluency of speech. By contrast, the term 'dysfluency' is used to refer to an abnormal disruption to fluency, such as following brain damage. The use of the prefix 'dys-' signals 'abnormal'.
Clause is a part of a sentence containing a subject and verb.

Some disfluencies may facilitate comprehension. For example, Fox Tree (2001) found that hearing an 'uh' aided listeners' recognition of subsequent words, suggesting that some disfluencies may act as cues that direct listeners' attention towards a particular word (see Box 13.1 in the next chapter). Hesitations have also been studied as potential cues to deception, as is explored in Box 12.2.

Box 12.2 Practical Application: Using verbal cues to detect a liar

There is a widespread misconception that it is easy to tell a liar from his or her demeanour, that certain behavioural cues reliably indicate when someone is lying. In fact, the average person cannot detect lies reliably and attends to the wrong cues when attempting to do so. Even people who might be expected to have had considerable exposure to deception are not necessarily effective at lie detection. Ekman and O'Sullivan (1991) examined lie detection by a sample of 509 people, including members of the US Secret Service, CIA, police and judges, as well as college students and other adults. They showed participants a videotape of 10 people who had been instructed to lie or tell the truth. Only the Secret Service personnel could detect lies at an above chance level, with average accuracy in this group of 64 per cent.

What can psychology tell us about how to detect lying? Vrij et al. (2001) note that someone who is lying, compared to someone telling the truth, tends to have a higher voice pitch, produces more hesitations and speech errors, speaks more slowly and uses fewer illustrators (e.g. hand gestures), and shows decreased hand, leg and foot movements. Furthermore, the liar's account of events produces fewer details than a true account (Vrij, 2004). These features are associated with the increased cognitive load involved in lying and are also observed when people engage in a cognitively complex or challenging task. People *perceive* lying to be associated with the following features: higher voice pitch, slow speech, more hesitations and errors, delayed responses to questions, averting eye gaze, smiling more often, and increased movements (Vrij, 2000). We therefore seem to attend to the wrong cues when attempting to detect lying, while people who attend to speech-related factors rather than behavioural cues would seem to be more effective at lie detection (Mann et al., 2004; Vrij & Mann, 2001).

Many studies of lie detection involve laboratory manipulations, with 'liars' instructed to deceive under various conditions. In real-life, lying may present differently, particularly when the stakes are high and the consequences of getting caught out in a lie are serious. Vrij and Mann (2001) examined the behaviour of a murderer who had been interviewed by the police on several occasions. The series of interviews had been videotaped. While the man initially denied that he was involved in the victim's death, the evidence suggested otherwise. He subsequently confessed to the crime and was convicted of murder. The videotaped interviews therefore allowed the researchers to examine the man's behaviour while lying (pre-confession) and telling the truth (during his confession).

While lying, the man produced longer pauses, slower speech and had more speech disturbances than when telling the truth, features that are consistent with lying carrying a heavier cognitive load. In a second part to Vrij and Mann's study, 65 police officers watched video fragments that had been selected from the interviews. The overall accuracy rate in the experiment was 64 per cent, which was significantly above chance; however, this resulted from good detection of truth rather than accurate detection of lies. In fact, accuracy of lie detection (at 57 per cent) was not above chance. The individual differences in lie detection were striking, and may reflect differences in selection of cues. Mann et al. (2004) found that police officers' accuracy was negatively correlated with reliance on stereotypical but non-diagnostic cues such as gaze aversion and fidgeting.

Such knowledge is now being applied to improve interview techniques. Vrij et al. (2010) support an 'information gathering' approach to interviewing that contrasts with the 'accusatory' approach often adopted. They note the advantages of the information gathering approach over an accusatory style: it increases the amount of detail provided by a suspect and is associated with more non-verbal and verbal cues to deceit being produced; it is less likely to elicit false confessions; and it is associated with lower interviewer confidence, which safeguards against bias. In addition, Vrij and colleagues suggest asking unanticipated questions to reduce the effects of preparation on the part of the suspect and asking questions about the precise time of events if a suspect appears to be using a scripted answer. Finally, because lying is more cognitively demanding than telling the truth, increasing the cognitive load may serve to differentiate liars from truths tellers. For example, having suspects give their account of events in reverse order increases cognitive load, as does asking event-irrelevant questions.

While speech does not provide us with a linguistic version of the 'Pinocchio's nose' with which to identify a liar, it does provide some relatively diagnostic cues that might be attended to. Such an approach would seem to be more effective than relying on stereotypical body language cues.

SLIPS OF THE TONGUE

Parapraxes are slips of the tongue or other actions originally thought to reflect unconscious motives.

In *The Psychopathology of Everyday Life* (1924/1938/1975), Freud treated speech errors as a particular class of **parapraxes** (action slips). While his emphasis was on supposed underlying repressed thoughts, he recognized that errors could be informative as regards language processing, asking 'whether the mechanisms of this disturbance cannot also suggest the probable laws of the formation of speech' (p. 71). So-called 'Freudian slips' are errors based on a substitution of a semantically or phonologically similar word (see Table 12.4), and, most researchers would now acknowledge, they reflect the cognitive processes underlying sentence formulation, rather than unconscious motivations or conflicts (e.g. see Norman, 1981; Reason, 1990, 2000).

Fromkin's (1971) analysis provided the first systematic account of error types. Fromkin showed that when errors occur they are not random; in fact they are systematic and are highly informative as to the nature of the underlying processing. The majority of speech errors are sound based errors (Fromkin, 1971, 2004) and errors tend to occur at one linguistic level (e.g. affecting phonemes or morphemes). Types of error are summarized in Table 12.4.

Table 12.4 Examples of types of speech error

Type	Description	Example(s)
Anticipation	Substitutions of a sound in anticipation of a sound that occurs later in the phrase. A full word can also be produced too early within a sentence	Cuff of coffee [cup of coffee]
Perseveration	The repetition of a sound from a previous part of the utterance	proliperation [proliferation]
Transposition/Exchange errors (also called metatheses or spoonerisms)	Transposition of two segments. Exchange errors can also affect words, where two words swap places in the sentence	You hissed all my mystery lectures [missed all my history lectures] You have tasted a whole worm [wasted a whole term]
Blend	A non-word is made based on two semantically related words	Mownly [mainly/mostly] Swinged [switched/changes]
Additions	A sound is added in	Similarily [similarly]
Deletions/omissions	A sound is omitted	Slit second [split second]
Semantic substitutions including Freudian slips	Retrieval of an incorrect but semantically related target	This room is too hot [cold]
Phonological substitutions or malapropisms including Freudian slips	A phonologically similar word is selected in error. Mixed errors, in which the target word and error share both semantic and phonological features, can also occur.	Projects [products] There's a pest in every class [pet] (this could be a deletion)

Lexical bias refers to tendency for phonological speech errors to result in real words.

Analysis of speech errors points to the importance of the phrase as a unit of production, as errors rarely jump across phrase boundaries. The vast majority of morpheme exchanges occur within clauses (Garrett, 1975). Errors preserve the consonant-vowel distinction and phonological errors are in keeping with the phonological constraints of a given language (Fromkin, 1971). Exchange errors show a **lexical bias** in that they are more likely to result in a word than a non-word (Rapp & Goldrick, 2000) – for example 'barn door' becoming 'darn bore' (Nooteboom & Quene, 2008). The frequency of lexical bias has been disputed; some researchers have argued that lexical bias is not common in spontaneous speech and is more likely to be induced experimentally (e.g. Garrett, 1980), but others argue that it also applies to natural speech (Dell & Reich, 1981). It would seem that the lexical bias effect reflects both immediate feedback between speech sounds and word forms (Dell, 1986) and monitoring of inner speech producing a real word bias (Levelt et al., 1999). In other words, non-word errors are more readily detected and repaired, while real word errors can 'slip through the net' and remain undetected before being uttered (Nooteboom, 2010).

Content words tend to exchange with content words and function words with other function words. Harley (2008) found no instances of content words and function words exchanging in his corpus of several thousand speech errors. Function words and bound morphemes (such as inflections) are generally left in place when a content word or morpheme moves, a pattern referred to as morpheme stranding. The following examples from Fromkin (1971) illustrate:

- nerve of a vergious breakdown [verge of a nervous breakdown];
- a weekend for maniacs [maniac for weekends].

Boomer and Laver (1968) found that stressed and unstressed syllables did not exchange with each other; errors were consistent with the stress pattern in the utterance. Furthermore, transpositions generally stay within the same syntactic or morphological class (Fromkin et al., 2010). These patterns of error suggest a systematic process whereby a sentence is constructed such that 'the word's skeleton or frame and its segmental content are independently generated' (Levelt, 1992, p. 10) and shows that speech production is highly rule governed (Fromkin et al., 2010).

Transposition errors, or metatheses, have been noted for centuries, and have become associated, perhaps unfairly, with the Reverend William Archibald Spooner, who was warden of New College, Oxford in the early 1900s. Many transposition errors were attributed to Spooner and this class of error has therefore become known as 'Spoonerisms'. Among the errors attributed to the Reverend Spooner (see also Table 12.4) are the following:

- you were fighting a liar in the quadrangle/lighting a fire;
- work is the curse of the drinking class/drink is the curse of the working class;
- the queer old dean/the dear old queen.

Many of these are apocryphal or exaggerated; if they were as frequently produced as some commentators have supposed it would suggest an underlying pathology (Potter, 1980).

Speech errors are generally collected from spontaneous speech, either by recording speech or having participants note their own errors in diaries, and large collections of speech errors can be accumulated in this way (e.g. Fromkin, 1971; Harley, 2008). An early example of such a corpus is provided by Meringer and Meyer's (1895) analysis based on a collection of an estimated 8,800 errors in German. Errors can also be induced experimentally, with various techniques developed to elicit errors. For example, the SLIP (Spoonerisms of Laboratory-Induced Predisposition) technique introduced by Baars and Motley (1974) aims to force transposition errors in research participants within a controlled setting. Another type of speech production error that has been induced experimentally with some success is the tip-of-the-tongue (TOT) state.

THE TIP OF THE TONGUE STATE

The **tip of the tongue** (TOT) state is a temporary inability to access a word from memory; when we experience a TOT, we generally can say whether we know the word and we may have access to some information about the word, such as its initial letter, what it sounds like, or whether it is a long or short word. In languages in which the noun has gender, that information can be available in the TOT state, showing that access to syntactic category information is preserved, although access to the specific phonological word form is unavailable (e.g. Vigliocco et al., 1997). This has implications for models of speech production, as is discussed below.

Tip-of-the-tongue (TOT) state refers to a temporary inability to access a known word.
Feeling-of-knowing is a subjective sense of knowing that we know a word, and is an example of meta-memory – our knowledge about the contents of our memories.

In the TOT state the target word is known to us, but we cannot access it; we have a **feeling-of-knowing** about the target, that is, we are aware that we know the word, yet we cannot produce it. In a review of research on the TOT state, Brown (1991) notes that it:

- is universal;
- occurs about once a week;
- increases in frequency with age;
- frequently affects recall of proper names;
- often involves an available initial letter;
- is often accompanied by other words;
- is resolved on almost half of occasions.

Brown and McNeill (1966) were the first to induce the TOT state experimentally. They read participants' definitions of rare words and asked them to produce the word to which they were referring. For example, they read the definition 'a navigational instrument used in measuring angular distances, especially the altitude of sun, moon and stars at sea'. The target word in this case was 'sextant'. The inability to access low frequency targets induced a TOT state in some participants (nine of 56 participants for this particular definition). Some of the incorrect words produced by participants show that some information is available and that lexical retrieval can involve partial activation. For example, in the case of the target *sextant*, some of the errors included *sexton*, *sextet*, *compass* and *protractor*; these are phonologically or semantically related targets. Bilinguals would seem to be more prone to the TOT experience (e.g. Gollan & Acenas, 2004), suggesting that competing lexical activations increase the likelihood of a TOT state occurring. However, the small differences in reaction times observed in the laboratory do not carry profound implications for everyday functioning of bilinguals (Baker, 2006) and should not detract from bilinguals' advantages in other domains (see Box 12.3).

Box 12.3 Research Close Up: Bilingual lexical access

Source: Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, *13*, 231–252.

INTRODUCTION

The cognitive effects of bilingualism have been studied for many years and a number of processing costs have been identified. Bilinguals would seem to be at a disadvantage compared to monolinguals on tasks involving the rapid retrieval of words from memory, though ability to access semantic information about those words is not impaired (Bialystok et al., 2009). For example, there is a bilingual disadvantage on the verbal fluency task, in which participants must name in one minute as many words as they can beginning with a given letter (e.g. name words beginning with the letter ‘s’) or belonging to a given category (e.g. names of animals).

METHOD

Sandoval et al. (2010) examined the verbal fluency of 30 monolinguals and 30 bilinguals when naming 15 semantic categories (e.g. ‘types of clothing’ ‘supermarket items’, ‘spices’) and 24 double-letter categories (e.g. words beginning with ‘fa’), with order of block presentation counterbalanced. The bilingual participants were English dominant English-Spanish bilinguals. Participants were instructed to name as many examples as they could think of that belonged to each category, with each trial allowing one minute for responses. In Experiment 1 responses were given in English; in a second experiment, responses were made in both languages.

RESULTS

The bilingual participants produced significantly fewer correct responses compared to the monolinguals in both the semantic and letter categories. While both groups of speakers performed better in the semantic category compared to the letter category, the bilingual disadvantage was equal for both types of stimuli. Furthermore, in the semantic task, the monolinguals produced responses of higher word frequency. Bilinguals were also slower than monolinguals to produce their first response at the beginning of a trial. Sandoval et al. suggest that interference between languages accounts for these differences. In a second experiment, responses were made in both languages. The bilinguals produced fewer responses in Spanish (the non-dominant language) than in English, and they were slower to produce their first response for a category in Spanish compared to English. Comparing performance in the two languages, Sandoval et al. found that bilinguals produced more cross-language intrusion errors when using their non-dominant language, while very few intrusions occurred when they spoke in the dominant language.

DISCUSSION

These data suggest that between-language interference in bilinguals creates processing costs that negatively affect verbal fluency, at least in the context of such experimental tasks. Does that mean that bilingualism is disadvantageous in terms of cognitive performance? The balance of evidence would suggest that the advantages of bilingualism far outweigh any potential disadvantages. In addition to the cultural, social and communicative advantages associated with bilingualism and multilingualism, cognitive advantages include greater flexibility and creativity in thinking (Baker & Prys Jones, 1998) and better executive function (Bialystok et al., 2008).

Several of the most influential theories of speech production have been based on analysis of speech errors. Here, we look at three models of speech production: Garrett's and Dell's accounts, which are based on speech error data and Levelt's account which takes a different approach.

THEORIES OF SPEECH PRODUCTION

It is generally agreed that there are a number of stages to speech production (e.g. Levelt, 1989). The first, conceptualization, is a poorly understood process by which a thought forms and is prepared to be conveyed through language. The processes by which an abstract thought or idea becomes a verbal thought remain elusive (think, for example, of the stages before you say to yourself 'I wonder if I turned the oven off before leaving the house'). The second stage involves the formulation of a linguistic plan. The concept or proposition must be translated so that the thought becomes language and the sentence that we want to output is planned. This process of translating from concept to language also remains a mysterious process, and if the goal of such theories is, as Clark and Clark (1977, p. 10) suggested, 'to discover how speakers turn ideas into words' we are arguably no closer to the holy grail. Levelt (e.g. 1989) considers the formulation stage as comprising two sub-stages. During the lexicalization sub-stage, the words are selected from the mental lexicon. In order for this to occur the concept must connect with the abstract word form or **lemma**. The lemma contains semantic and syntactic information about the target word but does not yet specify its phonological form. Formulation also involves syntactic planning; during this sub-stage the order in which the selected words will be output is decided. The third stage involves articulation of the plan. During this stage the sounds for the word are accessed, the **lexeme** is specified giving the full phonological form of the word, and the motor program for speech output is planned and articulated. In a final fourth stage, the output is monitored so that corrections can be made if necessary.

Lemma is an abstract word form that contains syntactic and semantic information about the word.
Lexeme is the basic lexical unit that gives the word's morpho-phonological properties.

The theoretical approaches to understanding speech production have much in common. It is recognized that clauses seem to be an important structure for speech planning and that processing proceeds from the abstract concept to syntactical processing, to precise phonological patterns. The various models of speech production differ in terms of the emphasis placed on these various components and in the extent to which they consider the processing stages to involve serial, parallel or interactive processing, that is, whether they favour a modular or interactive view of speech processing. Modular or serial theories posit a series of non-interacting stages, with different types of processing being completed at each stage (e.g. Fromkin, 1973; Garrett, 1980, 1982; Levelt et al., 1991a). On the other hand, interactive or parallel theories (e.g. Dell, 1986; MacKay, 1987) propose a less constrained account, with multiple sources of information operating to influence speech output. The debate over which type of theory provides the more accurate account continues.

MODULAR THEORIES OF SPEECH PRODUCTION

Garrett's model

Serial or modular theories propose that speech production progresses through a series of stages or levels, with different types of processing being completed at each level. According to Garrett's hierarchical model (e.g. 1980, 1982, 1992), speech is produced via a series of stages, proceeding in a top-down manner (see Chapter 2) so that processing

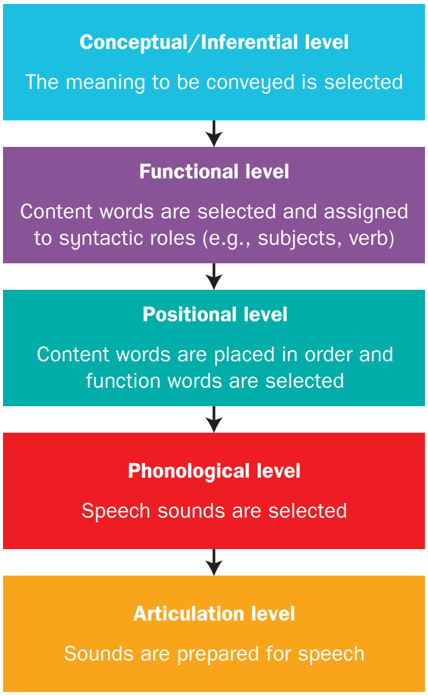

Figure 12.2 The stages of speech production proposed by Garrett (e.g. 1975). At the inferential level, we conceptualize the message that we want to express. At the functional level, the syntactic and semantic framework of the sentence is constructed. At the phonological level the sounds for the content words are acquired.

at lower levels does not influence that at higher levels (see Figure 12.2). Garrett developed his model to address patterns of errors in speech production (see also Dell, 1986, described below).

Figure 12.2 shows the various levels proposed by Garrett. At the inferential level, we conceptualize the message that we want to express. As noted previously, little is known about processing at this level or how the initial leap from thought to language-like representation might occur, particularly given that there may be a number of ways to express a thought. At the functional level, the syntactic and semantic framework of the sentence is constructed. At this stage, word exchanges would be predicted because the structure is present but the selected words, while activated, have not yet been allocated to their places within that structure or framework. Because the syntactic category of the word has already been determined, this model predicts that errors will not cross syntactic category (e.g. that nouns will swap with nouns but not nouns with verbs; verbs may swap with other verbs, though in reality verb exchanges are relatively rare). At the positional level, the words are allocated to positions within the syntactic frame. The function words are in place at this stage and so, where errors occur, bound morphemes tend to remain in their correct place (e.g. verb endings such as –ed and –ing). At the phonological level the sounds for the content words are acquired and sound errors can occur at this stage. In this model, the lexical bias effect described earlier in this chapter occurs during a later monitoring or editing stage where non-word errors are detected – word errors are more likely to 'slip through' undetected. The interactive models provide a rather different account of lexical bias, as we will see.

Evaluation

Garrett's model suggests that content and function words are treated differently and this is supported by the data on errors. One might argue that the relative sparing of function words reflects their higher frequency in language use (e.g. see Stemberger, 1985); however, bound morphemes have a lower frequency of use and yet are treated like function words in that they are retained in the syntactic frame of the sentence.

Non-plan internal errors occur when the intrusion is external to the planned content of the utterance.

Garrett's model provides a good account of the speech error data. However, Garrett's stages operate independently of each other and therefore this model does not predict errors that occur 'across levels'. For example, a type of error called a **non-plan internal error** occurs when concepts from the message level intrude when articulating the sentence, specifying the words at the phonological level. For example, someone intending to say 'let's get a coffee' while standing outside the library might say 'let's get a book'. Some 'Freudian slips' fall into this category, as a suppressed thought might interfere with current output; by Freud's account this was always the case: 'A suppression of a previous intention to say something is the indispensable condition for the occurrence of a slip of the tongue' (Freud, 1922, p. 52). Such errors have led to more interactive accounts of the stages of processing, such as that of Dell and colleagues outlined below. Before we look at interactive models, however, another influential modular account is considered, Levelt's model.

Levelt's model

Levelt and colleagues (e.g. Levelt, 1989; Bock & Levelt, 1994; Levelt et al., 1999) have presented a number of computational models of speech production leading to the sequential system called Weaver++ (Weaver stands for Word-form Encoding by Activation and Verification). This model focuses on the production of single words rather than the construction and output of whole sentences – it considers, for example, how we access (and say) the word 'cat' when we see a picture of a cat. A series of stages follow sequentially, from conceptualization to articulation (see Figure 12.3). Levelt's theory is based mainly on latency data (e.g. reaction times to picture naming) rather than error patterns, in contrast to Dell's and Garrett's models.

The theory focuses on the lexical access aspects of speech production. In Weaver++, the first two stages of processing involve lexical selection. Three stages of form encoding follow, before articulation occurs. These two systems, for lexical selection and for form encoding, would seem to involve quite different processes and involve different areas of the brain (Levelt, 2001).

The first stage is conceptual preparation, which Levelt et al. (1999) define as 'the process leading up to the activation of a lexical concept' (p. 3). The second stage involves lexical selection: a lemma or abstract word is retrieved from the mental lexicon and its syntactic category is activated. A number of words might be primed based on meaning, with selection dependent on relative activation so that the more appropriate selection occurs. For example, if I am shown a picture of a horse and asked to name it, the concepts *horse* and *animal* might be activated. These concepts activate the corresponding lexical items in the lexicon. This is an abstract word or lemma, which is 'essentially the lexical item's syntactic description' (Levelt, 2001, p. 13464).

Levelt et al.'s third stage involves morphological encoding. Once the lemma is selected, processing proceeds from the conceptual/syntactic domain to the phonological/articulatory domain; Levelt et al. recognize this as a crucial change. At this point, a TOT state can be produced – a lemma has been activated but the specific phonological form (lexeme) is not yet available. Because the lemma is a syntactic word, information about syntax is available, while the sound of the word is not yet accessed. This predicts the finding that in a TOT state a speaker has access to syntactic information, although they cannot produce the word. In many languages nouns have grammatical gender; this is unrelated to word meaning, that is, grammatical gender is a linguistic property unrelated to the conceptual properties of the referent. For example, in French *mouton* (sheep) is a masculine noun while *chèvre* (goat) is feminine. The word for 'milk' is masculine in French and Italian, and feminine in German, Dutch and Spanish. In languages with grammatical gender, information relating to noun gender can be activated in the TOT state, when the word

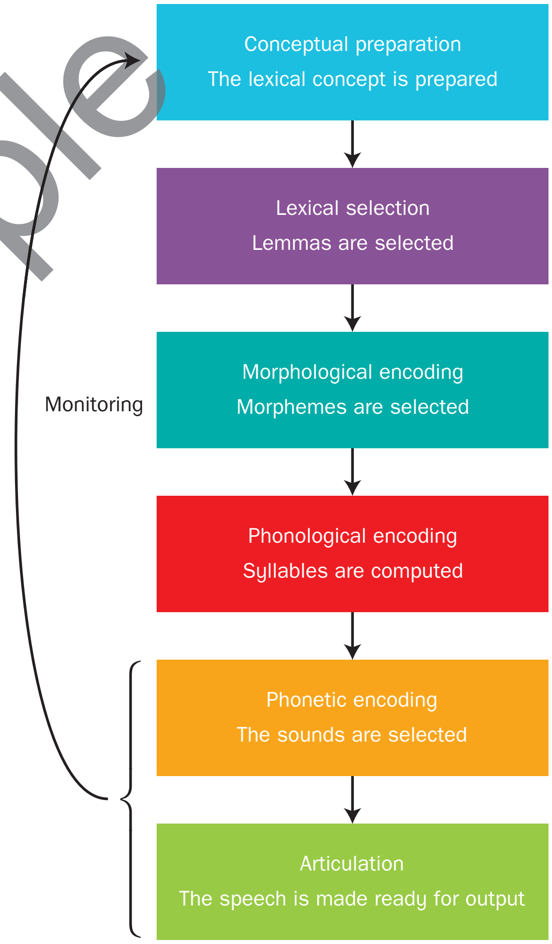

Figure 12.3 Levelt's model of speech production. In parallel to the processing stages, output monitoring allows the speaker to detect and correct errors.

Source: Adapted from Levelt, W. J. M., Roelofs, A. P. A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–37.

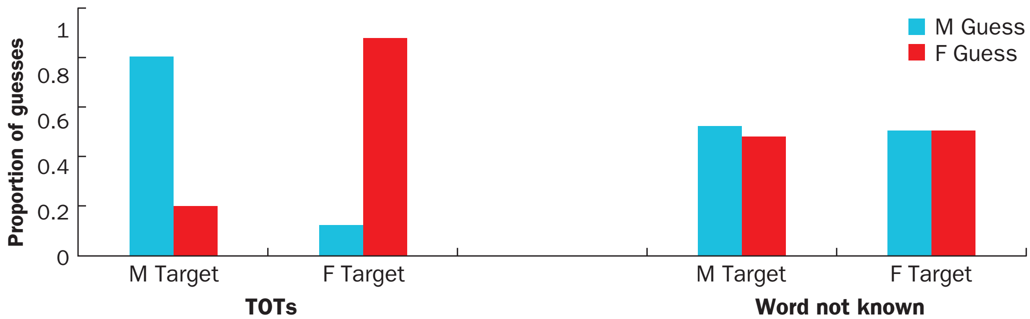

Figure 12.4 Responses to masculine and feminine target words for TOT words and words that could not be retrieved. Distribution of masculine (M guess) and feminine (F guess) responses for masculine (M target) and feminine (F Target) words, for TOTs compared to words that were not retrieved (here labelled 'word not known'), show access to knowledge of word gender for TOT words.

Source: Data from Vigliocco et al. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, *8*, 314–317.

itself is not accessible. Vigliocco et al. (1997) demonstrated this in Italian. Their participants were presented with definitions and asked to provide the corresponding word. Whenever a participant was unable to provide the word, they were asked to guess the gender of the noun (masculine/feminine), guess the number of syllables, give as many letters in the word as possible and state their position in the word, and report any other word that came to mind. Participants were subsequently shown the target word and asked whether it was the word they were thinking of. Vigliocco et al. found that noun gender was correctly reported 84 per cent of the time when participants were experiencing a TOT state (see Figure 12.4). By contrast, when participants could not produce the word and later could not affirm that the provided word was the target, performance was at chance level (53 per cent).

However, sometimes Italian speakers can access phonological information without accessing syntactic information (e.g. Miozzo & Caramazza, 1997), which does not support the notion of two separate stages for syntactic and phonological information – the distinction between semantic and phonological information remains. Processing at this stage involves three types of information; the word's morphology, its stress patterns, and the segments that make up the word are activated.

At stage four the syllables that make up the word are computed. Stage five performs phonetic encoding: the actual speech sounds activate at this stage. Levelt et al. (1999) posit the use of a syllabary, with phonological information allowing word articulation derived from the retrieval of syllables. These are 'highly overlearned gestural patterns, which need not be recomputed time and again. Rather they are ready made in the speaker's syllabary' (p. 5).

The sixth and final stage is articulation. The phonological information is transferred to a motor plan and executed by the articulatory system and speech musculature. The stages are presented in Figure 12.3.

The model considers the role of self-monitoring at multiple levels throughout the processing stages (Figure 12.3). This provides a mechanism for the detection of errors that allows us to repair speech and involves the cognitive mechanisms involved in speech comprehension (see Chapter 13). The precise means by which the mechanism operates and the attentional systems governing it are not elaborated.

Evaluation

Levelt's account shows how a series of specialized modules contribute to the process of speech production. It accounts for much data on speech production, including some patterns in bilingual speech. It also shows how speech might be monitored so that errors can be corrected. However, as a modular account, feedback between levels is limited. The retrieval of the word form occurs only after the lemma has been selected; there is no feedback from word form to lemma levels. However, some types of speech errors suggest that feedback does occur. Sometimes, the target word and the error share both form and meaning information, for example, saying 'rat' when you meant 'cat' (Treiman et al., 2003). This suggests that there is interference from lower to higher levels; Levelt et al. explain such errors as resulting from a failure in the monitoring processes. However, interactive models account for these and some other types of speech errors more successfully. We now turn to one influential account of an interactive type.

INTERACTIVE THEORIES OF SPEECH PRODUCTION

Dell's model

The final model of speech production that we will consider is Dell's cascaded or spreading activation account (Dell, 1986, 1995; Dell & O'Seaghdha, 1991; Dell et al., 1997), which is based on connectionist principles (see Chapter 1). This model uses the concept of spreading activation in a lexical network to show how competing activation across different levels might predict speech errors. In this model, activation from one level can affect processing at other levels, that is, processing is interactive. Processing is also parallel such that information can be processed at the different levels at the same time. These features, interactive and parallel processing, differ from the features of the serial models such as Levelt's.

There are four levels in Dell's model with processes corresponding to those described for Garrett's model above: that is, a semantic level, a syntactic level, a morphological (word) level and a phonological (sound) level. The semantic level is where we conceptualize what it is we want to say; at the syntactic level the structure of the sentence is devised; at the morphological level the morphemes that make up the target words are selected and at the phonological level the sounds that make up those words are activated.

Figure 12.5 illustrates the levels and connections. The connections between the layers allow bidirectional spreading of activation. That is, a word unit can activate the phonological units at the layer below (top-down spreading) and the semantic units at the layer above (bottom-up spreading; Dell & O'Seaghdha, 1991). According to this model, lexical access involves six steps:

1. The semantic units are activated by an external source (e.g. information from vision in a picture naming task, when you see a picture of the concept to be lexicalized or translated into words).
2. Activation spreads through the network.
3. The word unit with the highest level of activation is selected and linked to the syntactic frame for the sentence, in the appropriate slot. Once the word has been placed in the frame, its activation reduces to zero.
4. When the time is right, based on the slot in the syntactic frame the word is assigned to, the phonological information activates. If a single word is to be produced (e.g. in a picture naming task), selection of the word triggers the phonological information.
5. Activation continues to spread, but phonological units linked to the selected word become more highly activated.

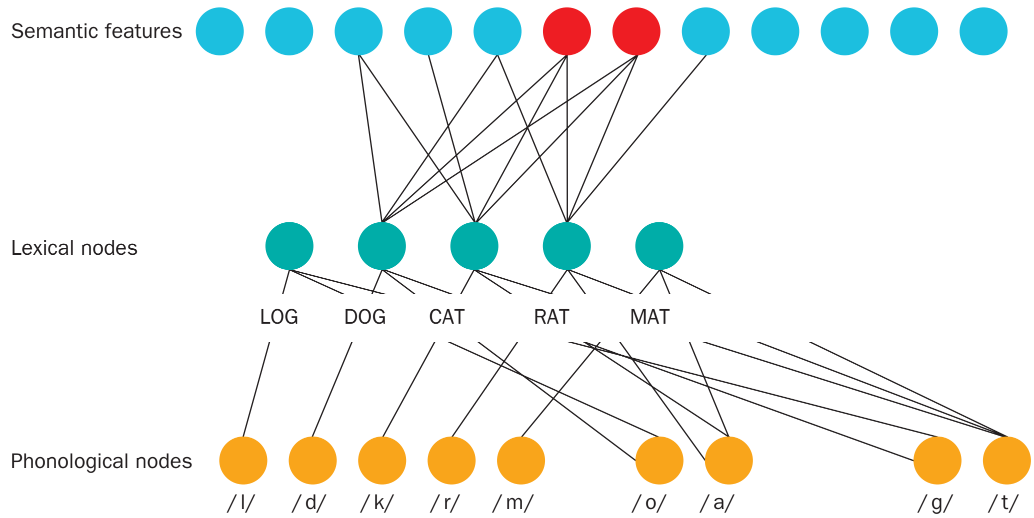

Figure 12.5 **Dell's spreading activation model.** The nodes in the top layer of Figure 12.5 represent the semantic features of the words. The words cat, rat and dog share semantic features and shared nodes are highlighted in red. The nodes in the middle layer represent lemmas or words. The nodes in the bottom layer represent the sounds that make up the start of the word, the vowel in the middle of the word, and the sound at the end of the word. Activation spreads throughout the network and the word that receives the most activation is the one that is selected (see also Levelt, 1999).

Source: Adapted from Dell, G. S., & O'Seaghdha, P. G. (1991) Mediated and convergent lexical priming in language production: A comment on Levelt et al. *Psychological Review*, 98, 604–614.

6 The most active phonological units are selected, and these are linked to slots in a phonological frame for the word; this allows the correct phoneme to drop into the correct 'slot' in the word so that the sounds are output in the correct order (see Dell & O'Seaghdha, 1991, pp. 605–606).

During the planning stage, the various words selected for the sentence become active; activation drops off once the word is placed in the sentence. This is an interactive rather than sequential account and feedback can occur from later to earlier levels such that phonological level activation can inform processing at earlier stages. Like Levelt's model, Dell includes a monitoring process to account for self-corrections and repairs. Errors occur when activation for a non-target overrides that of the target morpheme, phoneme or word. So word substitutions occur because a semantically related, but incorrect, choice achieves a higher activation than the target word. The lexical bias effect is accounted for by a backward activation process. Because words have lexical entries, they are represented by nodes in the network; non-words do not have associated nodes, but may have activation associated with phonotactic regularities. Words should therefore attain stronger activation and be more likely to be produced in error.

Similarly, the model accounts well for exchange errors (see Table 12.2 above). Exchanges of whole words (e.g. 'the sky in the sun') occur when a word of the same category – in this example, a noun – is dropped into the wrong slot in the syntactic frame. Once it has been output, its activation drops to zero, leaving the remaining noun to take the remaining noun slot in the sentence frame, because it has high activation and has not been selected.

Evaluation

This model accounts well for many patterns of speech error and some errors produced by people with aphasia are more in keeping with a parallel model than a serial model.

For example, Blanken et al. (2002) reported a mixing of word selection and word form access, in a German patient with aphasia, that supports an interactive account. Similarly, mixed errors (see above), in which the target word and the error share both form and meaning information, suggest that feedback does occur.

The spreading activation model deals with data from speech errors rather well. It also contributed to our understanding of sentence production rather than focusing on single word production alone, as is the case in Levelt's model. However, the model does not address the semantic level in any detail, focusing instead on the construction of syntactic, morphological, and phonological representations (Dell, 1986).

As yet there is no resolution to the debate between modular and interactive accounts of speech production, and a complete model may need to consider both modular and interactive aspects of the system. Dell and O'Seaghdha (1991) suggested that Levelt et al.'s (1991) data might be reconciled with spreading activation accounts by a 'characterization of the language production system as globally modular but locally interactive' (p. 604). The degree of informational encapsulation and interaction between components remains to be established in future research.

NEUROSCIENCE OF LANGUAGE PRODUCTION

Language involves a number of interacting brain areas, with many of the key language areas located within the left cerebral hemisphere in the majority of people. Language involves a number of cognitive processes, interacting with systems for attention, memory, perception and motor function. Sociocultural knowledge informs the ways in which we use language with others. Therefore many areas of the brain must contribute to language processing. **Neurolinguistics** is the study of the relationship between brain areas and language functioning.

> **Neurolinguistics** is the study of the relationship of brain function to language processing.

LATERALIZATION OF FUNCTION

Sensory information coming into one side of the body is processed on the contralateral (opposite) side of the brain. Similarly, fine motor movements such as hand movements are controlled by the contralateral cortical hemisphere. Information presented to the left visual field is processed in the right hemisphere, while information presented to the right visual field goes to the left hemisphere. The right side of the brain controls the left hand, and the left side of the brain controls the right hand. Information presented to each ear is processed in both hemispheres, but precedence is given to the contralateral side; stimuli presented to the right ear are processed in the left hemisphere.

Scan to watch a video on split brain experiments

Different functions are associated with the left and right cortical hemispheres. Language is largely a left hemisphere function while the right hemisphere is specialized for functions related to spatial/holistic processing (e.g. see Springer & Deutsch, 1981). When a cognitive function is *lateralized*, one cortical hemisphere is dominant for that function; for example, in most people the left hemisphere is dominant for language and the right hemisphere is dominant for face recognition. This **lateralization of function** is particularly apparent when we consider the effects on cognitive processing of a set of conditions that gives rise to the split-brain phenomenon. When the band of fibres connecting the two hemispheres, the corpus callosum, is severed, we can isolate the functions of the two hemispheres. In rare cases, these fibres may be severed surgically, to treat epilepsy for example. The corpus callosum may have failed to develop due to a developmental

> **Lateralization of function** refers to the asymmetric representation of cognitive function in the cerebral hemispheres of humans and higher primates.

condition. In such cases, the difference in the hemispheres' functions becomes more visible. As Sperry (a pioneering researcher in this field, and Nobel prize winner in 1981) (1974) noted (p. 7):

> Each hemisphere . . . has its own . . . private sensations, perceptions, thoughts and ideas all of which are cut off from the corresponding experiences in the opposite hemisphere. . . . In many respects each disconnected hemisphere appears to have a separate 'mind of its own'.

The split brain is explored further in Box 12.4.

Box 12.4 When Things Go Wrong: The split brain

We are unaware of the division of labour between the left and right hemispheres because the two hemispheres of our brains communicate so effectively via a number of connecting bands of fibres between the hemispheres, called commissures. The largest of these is the corpus callosum. This band of over 200 million fibres is surgically severed in a *commissurotomy*, a rare surgical procedure which is performed in order to reduce the effects of a type of intractable epilepsy that is unresponsive to drug therapy.

After the procedure the 'split-brain patient' behaves surprisingly normally, considering such a radical operation has been performed. However, on careful testing, it is apparent that the left and right hemispheres no longer communicate and are effectively working independently. The left hemisphere is dominant for language, in most people. If an object is placed in the right hand of a (blindfolded) split brain patient, he or she can name the object, as the information is relayed to the left hemisphere and it can make contact with the speech centre. However, if the object is placed in the left hand, the patient cannot name it. The patient can, however, pick a matching object from an array of objects, using the same hand. A picture that is presented to the right visual field can be named; a picture presented to the left visual field cannot, although, again, the object can be matched given an array of choices. Interestingly, when information is presented to the right hemisphere and cannot be named, the person reports not seeing it, suggesting a close alliance between language and subjective experience and consciousness (e.g. Cooney & Gazzaniga, 2003 ; Sperry & Gazzaniga, 1967). However, the patient can select a related picture, using the left hand, but, unaware of what the right brain saw, he or she will often invent a reason for the selection (see Figure 12.6).

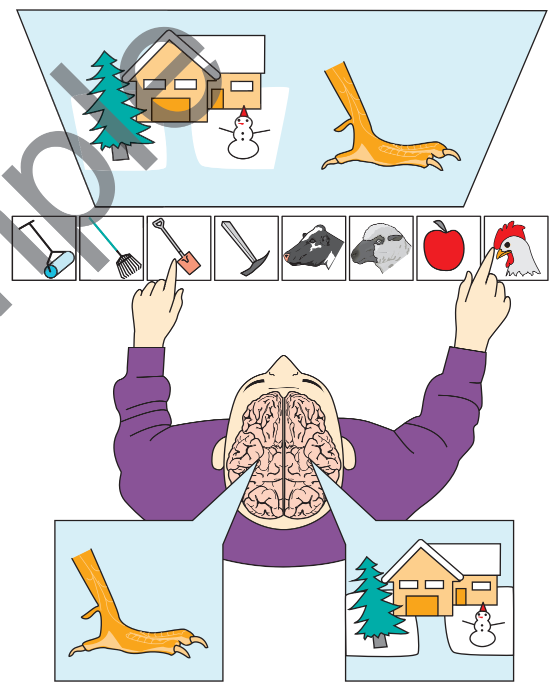

Figure 12.6 Demonstration of cognition in the split brain. A picture of a chicken claw is presented briefly to the right visual field, and goes to the (speaking) left hemisphere, while a snow scene is shown to the (non-speaking) right hemisphere. Asked to point out what he saw from a set of pictures, the patient's left hand points to a snow shovel while his right hand points to a chicken. When asked why he picked those particular pictures, the patient said 'Oh, that's simple. The chicken claw goes with the chicken, and you need a shovel to clean out the chicken shed.'

Source: Gazzaniga, M. S. et al. (Eds.) (1998). *Cognitive Neuroscience: The Biology of the Mind*. W.W. Norton & Co.

THE LEFT HEMISPHERE AND LANGUAGE

In the majority of people, speech is lateralized in the left hemisphere of the brain, and the left hemisphere is dominant for the majority of language functions. Rasmussen and Milner (1977), using the Wada test (a pre-surgical test of hemispheric dominance involving the selective anaesthesizing of the left and right hemispheres), found that 96 per cent of their patients who were right handed and over 70 per cent of those who were left handed had language lateralized in the left hemisphere (see also Kemp et al., 2008). Of the right handers, 4 per cent were right hemisphere dominant. Of the left handers, 15 per cent had bilateral representation of language, while 15 per cent were right hemisphere dominant. This dominance of the left hemisphere for language is evident in data from a number of sources: from studies of functional asymmetries in the typical population, from testing of split brain patients, and from the patterns of deficit seen in acquired language disorders such as aphasia. However, the degree of lateralization can vary in the typical population; there would seem to be, for example, differences in laterality between men and women (e.g. Shaywitz et al., 1995). Further consideration of differences in men's and women's language use is provided in Box 12.5.

Box 12.5 Research Close Up: Do men and women use language differently?

Sources: Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*, 581–592; Leaper, C. & Ayres, M. (2007). A meta-analytic review of moderators of gender differences in adults' talkativeness, affiliative, and assertive speech. *Personality and Social Psychology Review*, *11*, 328–363.

INTRODUCTION

An extremely lucrative section of the self-help industry has sprung up around the notion that men and women communicate in radically different ways, one bestselling book even relying on the metaphor that men and women 'come from different planets'. This focus on differences supports stereotypes of gender differences and ignores evidence to the contrary.

Differences in language and communication, when they are found, are surprisingly small, and the distribution of men and women on such measures mostly overlap, as two meta-analytic studies demonstrate.

METHOD AND RESULTS

Meta-analysis is a statistical technique that allows comparisons to be made across independent studies that can be combined on the basis that they address similar research questions: the effectiveness of a meta-analysis depends to a large degree on the criteria used to include studies in the analysis. Hyde (2005) reported a set of meta-analyses of research studies reporting gender differences across six categories, assessing cognitive variables such as abilities, verbal or non-verbal communication, personality variables, measures of psychological wellbeing, motor behaviours, and miscellaneous constructs such as moral reasoning. Hyde found that 78 per cent of the studies reporting gender differences showed small effect sizes (Cohen's d (a measure of the difference between two means) range of 0.00 to 0.35), with a median effect size of 0.21 (Fiske, 2010). Strong gender differences were supported only for motor skills such as throwing speed and distance, on some measures of sexuality (such as reported incidence of masturbation and attitude to sex outside of a committed relationship) and physical aggression. Measures of vocabulary, reading and speech production showed small or negligible effects, and any stronger effects are likely to be context dependent. For example, one of the stronger (though still small) effects among the communication variables suggested that men are more likely than women to interrupt a conversation (d = 0.33). An effect like this is likely to be influenced by other factors such as personality, topic of conversation and factors concerning the other interlocutors. Yet, the stereotype that men interrupt and women don't prevails.

Leaper and Ayres (2007) conducted meta-analyses of studies examining gender differences in adults' talkativeness, affiliative speech, and assertive speech. While there were statistically significant effects for all three language constructs, the average effect sizes were negligible. Contrary to the stereotype, men were found to be more talkative than women (d = 0.14) across the studies examined. Consistent with the stereotype, men used more assertive speech (d = 0.09), while women used more affiliative speech (d = 0.12). However, in all cases the effect seizes are so small that it would not be prudent to conclude that there are differences, particularly when there is inconsistency across studies examining these constructs. Studies of these language constructs in children have also produced very small effects (e.g. Leaper & Smith, 2004: talkativeness, d = 0.11 in favour of girls; assertive speech, d = 0.11 in favour of boys; affiliative speech, d = 0.26, in favour of girls).

DISCUSSION

To conclude that there are robust sex differences in speech production, we would expect to see repeated demonstrations of the difference across a variety of studies and contexts and always in the same direction. We would also expect to see consistent strong effect sizes, as occurs when we measure throwing distance, for example. This is quite simply not the case, and it would seem that gender differences in language ability have been greatly exaggerated. Wallentin's (2009) comprehensive review of sex differences in verbal abilities and language cortex concluded that there are no sex differences in language proficiency and highlighted the inconsistent findings relating to differences in language-related cortical areas. Wallentin also notes the problem of publication bias – studies that have found statistically significant differences are more likely to be published than those showing null findings. We therefore do not know how many unreported studies finding no gender differences may be out there.

EVIDENCE FROM THE TYPICAL POPULATION

Dichotic listening task is one where different stimuli are presented to each ear.

The **dichotic listening task** (see also the section on attention in Chapter 3) involves the simultaneous presentation of stimuli to the left and right ear. While auditory processing involves both contralateral (opposite side) and ipsilateral (same side) connections from ear to brain, contralateral connections are dominant; that is, while stimuli presented to the right ear are received by the auditory cortex of both cerebral hemispheres, the dominant connections are contralateral and therefore verbal stimuli presented to the right ear are predominantly processed by the left hemisphere (Kimura, 1967). Tests using the dichotic listening task have shown that there is a right-ear advantage for verbal stimuli (see Chapter 13). Participants report more words (or speech sounds) that have been presented to the right ear compared to the left (Springer & Deutsch, 1981). This left hemisphere specialization seems to be in place at quite a young age, as children as young as two years of age show a right-ear advantage for speech sounds (Hiscock, 1988) and infants under 10 months show greater left hemisphere activity when brain waves are measured during presentation of speech (Molfese & Betz, 1988). Studies suggest that the right-ear advantage may be restricted to consonant sounds (Best, 1988). Consonants and vowels may be treated differently; many non-speech vocalizations, and even the calls of apes, are vowel-like. On the other hand, the rapid changes in consonant sounds evident in human speech are complex auditory patterns and require high level sequential processing (some would argue that language is a special case of sequential processing).

Different areas within the left hemisphere process information relating to meaning and to syntax. Another way in which we can examine language in the normal brain is through measuring event-related potentials (ERPs; see Chapter 1). ERPs provide high temporal resolution (meaning that very quick changes in brain activity, with time scales of milliseconds, can be detected) and are tied to a particular event, such as a stimulus presented to a research participant. Electrodes are placed on the scalp to measure changes in electrical activity in the cortex as stimuli are presented. The changes in the brain waveform are informative as regards the nature of language processing. For example, Kutas and Hillyard (1980) compared brain waves as normal, semantically anomalous and physically anomalous sentences were presented (as written stimuli) to participants. Brain activity in response to semantically incongruous sentences differed from that seen when physically incongruous sentences were presented (see Chapter 13), supporting the notion that syntactic and semantic information are treated differently and processed in different areas of the brain (see also Hagoort & Brown, 2000; Osterhout & Holcomb, 1992).

Language has also been studied in the normal brain using a method called **transcranial magnetic stimulation** (TMS). TMS is a non-invasive way of stimulating particular cortical areas such that their functions are facilitated or inhibited. While the stimulation is short-lived and effects are largely temporary, in some circumstances TMS has been shown to alter the functioning of the brain beyond the initial period of stimulation, and therefore the method has implications in terms of therapeutic interventions. For example, Wirth et al. (2011) used TMS to enhance participants' performance on an overt picture naming task, and de Vries et al. (2010) used TMS to enhance grammar learning, or specifically the ability to detect syntactic violations. De Vries et al. showed that Broca's area plays a crucial role in grammar processing. Such studies suggest potential applications to remedial intervention in cases of language disorder after brain injury.

Transcranial magnetic stimulation is a non-invasive method of temporarily exciting or inhibiting cortical areas.

So far we have considered the left hemisphere as the site of language processing, but the right hemisphere also has a role to play, albeit a supporting role. The right hemisphere is involved in emotional aspects of speech, prosody and aspects of non-literal speech. People who have damage to the right side of the brain have difficulty in appreciating the emotional tone of an utterance (Caplan, 1987), and they have difficulty in understanding non-literal speech such as sarcasm, figurative language and indirect requests (Weylman et al., 1988), suggesting a role for the right hemisphere in processing the pragmatic aspects of an utterance. Similarly, studies of individuals with split-brain syndrome show that the right hemisphere is very limited when it comes to syntactic and phonological processing but it may be capable of other language functions, albeit not in the specialized way of the left hemisphere (e.g. Gazzaniga, 1983). While the right hemisphere is involved in the processing of tone in non-tonal languages (such as English), it is the left hemisphere that processes tone in tonal languages (such as Chinese) in which tone carries meaning (Gandour et al., 1992).

EVIDENCE FROM APHASIA

Speech production results from the processing that occurs in a number of language areas located around the Sylvian fissure of the left hemisphere, an area referred to as the peri-Sylvian language region. Damage to any of these areas can impair the ability to produce speech or writing. We can learn much about speech processing in normal cognition by examining the ways in which language is affected by brain injury.

Some of the key left-hemisphere language areas are shown in Figure 12.8; the figure does not reflect the considerable individual variability in the functional localization of

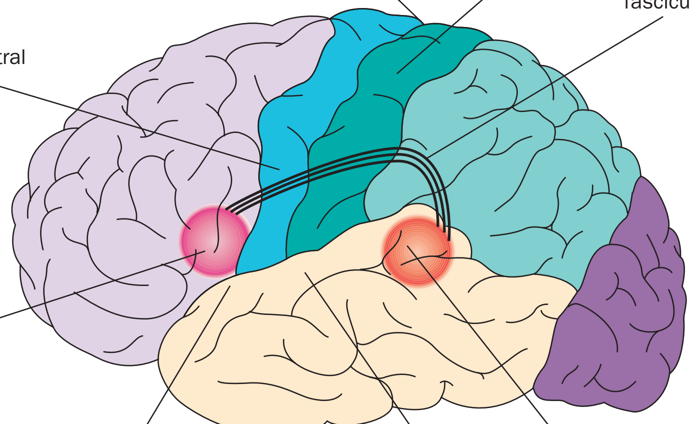

Figure 12.7 **Key left hemisphere language areas described by the Wernicke–Geschwind model.** In fact, the precise location, and role, of the language areas continue to be debated, not least because of the considerable individual variability that is evident in the functional localization of language.

Electrocortical stimulation of the surface of the cortex allows a surgeon to locate, and avoid damage to, brain regions associated with a particular cognitive function.

Wernicke-Geschwind model is a simplified model of language function used as the basis for classifying aphasia disorders.

language, however. In patients undergoing surgery to these areas, **electrocortical stimulation** while the patient is awake allows the surgeons to locate individual language areas and to reduce the risk of post-operative neurological deficits.

The **Wernicke-Geschwind model**, proposed by Karl Wernicke (1874), and later Norman Geschwind, notes a number of key areas for language (see Figure 12.7) and presents a simplified account of their role in language processing. The model proposes that we repeat a heard word by processing of the following sequence of brain areas. Following processing of the word in the auditory cortex, information about word meaning is processed in Wernicke's area and the output is sent to Broca's area via a band of connecting fibres called the arcuate fasciculus (see Figure 12.7). Broca's area prepares the speech output and a motor program for output is then articulated via the motor cortex. When we read a word out loud, a similar sequence is involved, with processing starting at the back of the brain in the primary visual cortex and continuing into Wernicke's area via the connections of the angular gyrus. While this model represents a simplification of the processing involved, it does provide a useful overview of the principal cortical brain areas for language and their functions.

Aphasia (or dysphasia as it is sometimes called) is the term used to describe a deficit in language following brain injury. It generally refers to spoken language, with the terms agraphia and alexia used specifically for deficits in writing and reading, respectively. In aphasia, the internal processing of language has broken down; it is not that the person's muscles or motor control for producing language have been damaged (as may occur in conditions such as apraxia and anarthria, for example). In a small percentage of people, damage to the right side can produce aphasia; aphasia following right hemisphere damage is called **crossed aphasia**.

Crossed aphasia refers to language dysfunction following right hemisphere damage in a right-handed individual.

Aphasic disorders can be classified according to whether they are of fluent, non-fluent or pure type. In the pure disorders a particular facet of language (e.g. the ability to repeat back sentences) is affected, while other language functions remain intact. The fluent disorders are characterized by fluent but empty speech, that is, the person produces fluent sentences, but the content of the utterances is not as they intended. The non-fluent disorders are characterized by reduced speech output, slow or effortful speech. Generally, damage to the anterior regions (near the region marked as Broca's area in Figure 12.7) creates a non-fluent type of disorder while more posterior damage (near Wernicke's) can cause a fluent type of aphasia. However, the site of damage can vary considerably from patient to patient, even with similar deficits, and younger people tend to show a non-fluent pattern of deficit regardless of the site of damage; fluent disorders are very rarely found in children, for example (e.g. Murdoch, 2009). In what follows, we will describe in a general way the main deficits associated with each category of aphasia; in fact, people with aphasia show quite a range of individual differences in performance on language tasks, and in terms of recovery of function.

Broca's aphasia

One of the first cortical areas involved in language production to be identified occupies the left inferior frontal gyrus and is known as Broca's area (see Figure 12.7). In 1861, a French doctor, Paul Broca, localized language to the left hemisphere, and attributed the production of speech to the area now named after him. (A paper by Marc Dax, dated to 1836, is now acknowledged as the first to identify the left hemisphere as the seat of language.) Broca's account was based on the aphasic disorder of a patient he encountered at the Bicêtre hospital in Paris. The man, called Leborgne, presented in his twenties with a severe reduction in speech output; over the subsequent years he gradually lost the use of his right arm and leg, an impairment confirming left hemisphere damage, as limb movement is largely controlled by the contralateral cortical hemisphere. Broca wrote of Leborgne (1861a):

Broca's aphasia is an acquired language disorder characterized by non-fluent speech, reduced speech output and problems with grammar processing.

Scan to see a patient exhibiting Broca's aphasia

> He could no longer produce but a single syllable, which he usually repeated twice in succession; regardless of the question asked him, he always responded: tan, tan, combined with varied expressive gestures. This is why, throughout the hospital, he is known only by the name Tan.

Broca called this disorder 'aphemia' (meaning 'without speech'), as he believed that the patient could understand language and therefore that the speech deficit was independent of language function itself; that it was a specific problem with voluntary speech. The similar term 'aphasia' was subsequently coined by Trousseau (Broca, 1864; cited in Dronkers et al., 2007). It is now recognized that there are some comprehension problems associated with Broca's aphasia, and these problems are particularly apparent when test sentences move beyond simple syntax (e.g. passive voice constructions). It is also now clear that it is the abstract representation of speech that is impaired in Broca's aphasia; in deaf signers with aphasia, the linguistic components of sign language are similarly affected (LeBrun & Leleux, 1986; Poizner et al., 1984).

After Leborgne's death at the age of 51 years, the area now known as Broca's area was discovered on autopsy to have been damaged by infection, which left a large abscess in that region. Broca concluded that this area of the brain was responsible for speech production. Leborgne's deficit was severe, leading some modern commentators to question whether he had a more **global aphasia**.

Global aphasia is an acquired language disorder involving extreme impairment of language function.

Non-fluent aphasia is when the patient's speech output is reduced, laboured, or absent.

Broca did not dissect the brain but preserved it and so his analysis of the damage was restricted to a surface inspection. Dronkers and colleagues (2007) were able to access Tan's preserved brain, and subject it to high resolution MRI scanning. They found substantial lesions extending into medial regions of the brain, in addition to the surface lesions that Broca reported, suggesting that global aphasia may be the more likely diagnosis.

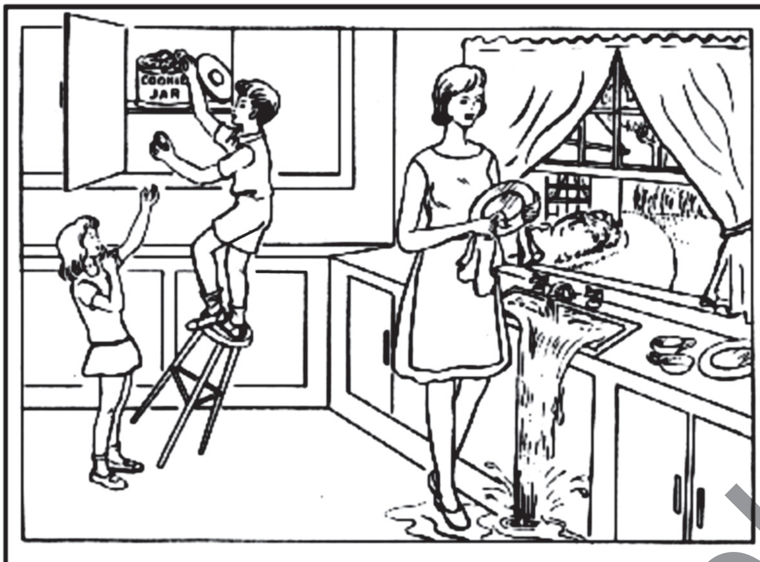

Figure 12.8 The cookie theft picture from the Boston Diagnostic Aphasia Examination. The picture shows a number of characters and actions within a familiar scene (a stereotypical kitchen scene) and elicits predictable spoken descriptions involving nouns (boy, girl, cookie, stool, silk, water), and verbs (looking, taking, spilling, falling) as well as discourse around intentions (such as to take without being seen).

Source: Cookie Theft picture. (Adapted) From the Boston Diagnostic Aphasia Examination – Third Edition, by Howard Goodglass in collaboration with E. Kaplan and B. Barresi, #11880, Austin, TX: PRO-ED. Copyright 2001 by PRO-ED, Inc. Reprinted (Adapted) with permission.

Patients with Broca's aphasia show deficits ranging from severe mutism to dysfluency or laboured speech. Broca's aphasia is one of a number of disorders that are categorized as **non-fluent**, expressive or productive aphasia. Speech output is reduced and non-fluent, but word selections tend to be meaningful. Function words (those that do the grammatical work in a sentence) rather than content words tend to be compromised. People with non-fluent aphasia are aware of their speech problems, which has implications for testing, as motivation may be an issue.

The following excerpt from Buckingham (1981) illustrates the marked dysfluency and reduced output associated with non-fluent aphasia. In this excerpt, B.L., a patient with Broca's aphasia, is trying to describe a picture from the Boston Diagnostic Aphasia Examination, called the cookie theft picture (see Figure 12.8).

The description is as follows:

> B.L.: Wife is dry dishes. Water down! Oh boy! Okay Awright. Okay . . . Cookie is down . . . fall, and girl, okay, girl . . . boy . . . um . . .
>
> Examiner: What is the boy doing?
>
> B.L.: Cookie is . . . um . . . catch
>
> Examiner: Who is getting the cookies?
>
> B.L.: Girl, girl
>
> Examiner: Who is about to fall down?
>
> B.L.: Boy . . . fall down!

A number of features of Broca's aphasia are apparent from this excerpt. First, reduced output is apparent. This type of speech output is sometimes known as telegraphic speech, because the sentences are reduced to the most basic units required to convey meaning – the content words such as nouns and verbs. The selection of content words is correct, showing that the patient can access the words from the lexicon. The function words are by comparison relatively sparse – inflections such as verb endings, conjunctions (e.g. and, but) and prepositions (to, under) are absent.

Goodglass and Geschwind (1976) defined Broca's aphasia as a condition 'marked by effortful, distorted articulation, reduced speech output, and agrammatic syntax but sparing of auditory comprehension' (p. 237). However, as mentioned above, while comprehension of everyday conversation may be relatively intact, people with Broca's aphasia have difficulties in understanding complex syntax. When comprehension depends on processing and understanding the syntactic structure of the sentence, it fails (Cornell et al., 1993).

Broca's area is just one part of the left inferior frontal cortex; a number of areas have been identified, within that region, that contribute to various aspects of language function. For example, separate areas for semantic and phonological processing have been identified within the left inferior frontal cortex (e.g. see Gough et al., 2005). Grodzinsky and Santi (2008) provide a useful overview of the state of current knowledge around the functions of Broca's area and reiterate a key role in syntactic processing (see also de Vries et al., 2010).

Wernicke's aphasia

Wernicke's aphasia is a fluent aphasia, characterized by fluent but meaningless output and repetition errors.
Fluent aphasia is when the patient's speech is fluent, but not meaningful.

A few years after Broca's discovery, Carl Wernicke reported a contrasting pattern in two patients who, after brain injury, showed normal pace and intonation but jargon-like speech. Wernicke's aphasia is associated with damage further back in the brain than the region associated with Broca's aphasia, in the upper part of the left temporal lobe (and extending to the angular gyrus, and the supramarginal gyrus; this region is the inferior parietal lobule). Wernicke's aphasia is classified as a **fluent aphasia**. While speech output is fluent, it is empty, that is, it is not meaningful. This condition was sometimes referred to as neologistic jargonaphasia, the word neologism referring to the patients' tendency to produce non-words, which may reflect partial activation of phonological information (Ellis et al., 1983). These patients are generally unaware of the problem with their speech output. An example from Goodglass (1983) illustrates some of the characteristic speech patterns of the disorder, again using the cookie theft picture (see Figure 12.8):

Scan to view a patient exhibiting Wernicke's aphasia

> Well this is . . . mother is away here working her work out o'here to get her better, but when she's looking, the two boys looking in other part. One their small tile into her time here. She's working another time because she's getting, too.

Another excerpt from Goodglass (1993, p. 86) shows similar features. In this case, the patient was responding to being asked 'How are you today?': 'I feel very well. My hearing, writing been doing well. Things that I couldn't hear from. In other words, I used to be able to work cigarettes I didn't know how . . . Chesterfeela, for twenty years I can write it.'

People with Wernicke's aphasia are likely to produce phonemic paraphasias, that is substitution errors in which a similar sounding word (or non-word) is produced instead of the target word (e.g. 'why' for 'wine'). There are made-up words or neologisms and overall there is a striking contrast to the pattern seen in Broca's aphasia. The function words and the grammatical structures of the sentences produced are relatively intact; the problem concerns the content words. Wernicke (1874) speculated that while Broca's area was involved in motor programs for speech output, the area now known as Wernicke's was involved in processing sounds for meaning. He also speculated as to what would happen if the connections between the two areas were severed: the patient would have difficulty repeating back what was said.

Table 12.5 Summary of language deficits in aphasia and site of damage

Type	Lesion site	Effect on speech output	Other deficits
Broca's aphasia	Anterior	Non-fluent output, reduced effortful speech	Repetition Naming
Wernicke's aphasia	Posterior	Fluent but 'empty' or meaningless speech	Comprehension Repetition Naming
Conduction aphasia	Arcuate fasciculus	Fluent	Repetition Naming
Anomic aphasia	Can be anywhere in language region	Fluent but with word finding difficulty	Naming
Global aphasia	Large area of damage	Extremely limited language function	Comprehension Repetition Naming

Conduction aphasia is when the patient has a specific difficulty affecting the repetition of speech.

The arcuate fasciculus was identified as the band of fibres that connects Broca's and Wernicke's areas (in fact, it is now known to link Wernicke's and the motor/premotor frontal areas) and 'disconnection' of this band of fibres is associated with a specific deficit in repetition, a disorder known as **conduction aphasia**. Bartha and Benke (2003) outline the main characteristics of conduction aphasia: severely impaired repetition, frequent phonemic paraphasias (saying unintended syllables or words, e.g. saying 'whine' instead of 'while'), repetitive self-corrections and word-finding difficulties. Repetition deficits are a key feature; spontaneous speech is generally fluent, although paraphasic, and comprehension is close to normal. Neuroimaging studies have revealed that the neurological bases of conduction aphasia are more complex than originally thought. As Ardila (2010) notes, relatively few cases of conduction aphasia have a lesion limited to the arcuate fasciculus, and, furthermore, conduction aphasia can occur when damage is limited to the cortex, without subcortical lesions. The main symptoms of the different categories of aphasia are summarized in Table 12.5.

Anomic aphasia

Anomic aphasia is when the patient has a specific difficulty with word retrieval.

Anomia is a word finding disorder that has been likened to the TOT effect in normal speech. Relatively small lesions within the language areas can produce anomia, as can transient conditions that reduce blood supply to these areas (Obler & Gjerlow, 1999). For the individual with anomia, access to the word that he or she is searching for is denied, but the patient has not lost knowledge of the word or of its meaning. Allport and Funnell (1981, p. 405) illustrate one patient's word finding problem with the following excerpt, again using the cookie theft picture in Figure 12.8 (the square brackets shows their guess as to what the patient was trying to say):

> Well it's a . . . [kitchen] it's a place and it's a girl and a boy and they've got obviously something which is made . . . some . . . [biscuits], some . . . made . . . well . . . [the stool] it's just beginning to . . . [fall] go and be rather unpleasant . . . and . . . this is the [mother?] the woman, and she is [pouring?] putting some . . . [water] stuff. . . .

The same patient could select the correct name for an object when shown a picture and two written object names as long as the two words were not related in meaning. Therefore it is not that knowledge of words is impaired; rather the patient's ability to access the words is deficient. Allport (1983) suggests that this reflects a problem with translation between word forms and their conceptual representations. A similar pattern has been observed in developmental disorders of language such as specific language impairment (SLI). Constable et al. (1994, p. 1) reproduced the speech of a seven-year-old boy with SLI as he tried to name a set of handcuffs presented in a picture naming task:

> Key . . . oh what do you call them . . . oh yeah . . . you put . . . you put . . . with your . . . with your . . . oh . . . with your . . . when you . . . when someone's stole something . . . and . . . what do you call them . . . necklace? . . . no . . . I just don't know the word.

Evaluation

Neuropsychological cases have contributed valuable data towards cognitive accounts of speech and language production. However, we must be cautious in interpreting data from such cases. As mentioned above, there is considerable variation between patients; people with the same pattern of deficit can have damage to different areas and those with similar damage can have differing language deficits. Lesions to Broca's area can occur without Broca's aphasia (Dronkers, 1996) and a Broca's-type aphasia can follow damage to areas outside Broca's area (Caplan & Hildebrandt, 1988). Furthermore, particular types of aphasia are generally associated with a reduction in a particular behaviour (use of function words, for example) rather than a complete absence of such features (Kolk, 2007). These factors must be taken into account when considering aphasia syndromes as applied to models of normal language use.

Box 12.6 Practical Application: Supporting language expression in 'locked-in syndrome'

In the book *The Diving Bell and the Butterfly* (1997), French journalist Jean-Dominique Bauby described his life after a stroke affecting the brain stem left him with a condition known as 'locked-in syndrome'. Patients with the condition are described as 'locked in' because, although they are awake, they cannot move; they are essentially locked inside their paralysed bodies. A particularly distressing aspect of the condition is the inability to communicate with others. As Gosseries et al. (2009) note, 'testimonies from victims relate that the worst aspect of the experience was the anxious desire to move or speak while being unable to do so' (p. 192).

In 'classic' locked-in syndrome, eye blinks and up-down eye movements remain intact, allowing some patients to communicate by means of an eye blink or eye movement system. Patients with total locked-in syndrome are unable to produce any voluntary movements (Gosseries et al., 2009; Kübler et al., 2001). The ingenuity of patients in such situations is remarkable; in one case, a patient used eye blinks to communicate via Morse code (Feldman, 1971).

Jean-Dominique Bauby painstakingly dictated his 136-page memoir using eye blinks. A frequency-ordered alphabet (e.g. in English: E-T-A-O-I-N-S-R-H-L-D-C-U-M-F-P-G-W-Y-B-V-K-X-J-Q-Z) was read aloud to him – he selected the target letter by blinking his left eyelid as the letter was read. The system relies on an effective partnership between the patient and an interlocutor, who notes the letters and may become adept at predicting the patient's intended word. With practice, the system can allow the patient to communicate immediate concerns, though it does not lend itself to conversation and the patient has to rely on the interlocutor to convey his or her meaning.

Advances in augmentative and alternative communication technologies have facilitated communication for those with locked-in syndrome. Eye-controlled computer-based technology reacts to patients' eye movements, allowing them to navigate a computer system with eye movement fixations functioning much as a mouse is used to move a cursor. The computer screen can show a keyboard, from which patients select letters by eye fixation. A speech synthesizer can then allow their completed sentence to be read out loud. Menu keys also allow the user to control the environment, for example, the patient could call for assistance, browse the internet or send an email (Gosseries et al., 2009). The advantage of this system is the control and independence it affords the patient. Recent developments in brain–computer interfaces (BCI) hold further promise. BCI systems make use of brain activity, via electroencephalogram (EEG) oscillations or event-related brain potentials (ERPs), or blood flow based measures, for example, to drive external devices in real-time, essentially allowing the impaired motor system to be bypassed (Birbaumer et al., 2008). While such systems are still in their infancy and remain to be fully tested in patient populations, the results so far are promising.

WRITING

Having looked at the processes involved in the production of speech, this chapter ends with a brief overview of another form of language production: writing. The processes involved in writing are similar to those involved in speech production, but writing requires access to the orthographic (written) form of a word rather than its phonological form. Writing processes will therefore differ to some extent according to the orthographic properties of the language used (see also the section on reading in Chapter 13). Writing also differs from speaking in that when we write we have more time to think about what it is we want to express and to 'translate' it into a written form. We can take our time over the construction of sentences, whereas speaking is time pressured. We can also monitor the output more easily; we can read the sentence we have written and inspect it, and correct it if necessary. Unlike speech, writing is often a solitary activity; while a writer will have a reader in mind when writing, he or she lacks the immediate feedback that occurs during a spoken conversation. Another difference between discourse and writing is that writing makes fewer demands on memory and therefore more complex ideas can develop through writing. The act of writing ideas down can facilitate thinking and bring about deeper understanding of the subject matter (Pijlaarsdam et al., 1996).

Research examining writing focuses on the later stages of the process, including composition and revision processes, as the earlier processes such as lexical retrieval and structuring a sentence have much in common with speech processes. Writing a textbook or an essay involves a number of higher cognitive processes collectively referred to as composition. Composition is a process by which ideas are turned into symbols (Kellogg, 1999). Like speech production, composition involves the various components of language, with smaller units contributing to the overall goal of discourse-level output. As is the case for speech production, writing involves a number of stages.

THE HAYES AND FLOWER MODEL OF WRITING

Hayes and Flower (1980) proposed a cognitive model of writing that focuses on three main domains affecting the writing process. These are the task environment, long-term memory and the immediate cognitive aspects of the writing process. The task environment includes the topic of focus, the intended readership and the purpose of the writing task. The writer must have an accurate understanding of these factors in order to progress the writing task. For example, if you are writing an essay on the psychology of language, you must identify the topic and the main points, write with a reader in mind (in this case an examiner who will be grading the essay) and consider the factors that will lead to the essay receiving a good grade.

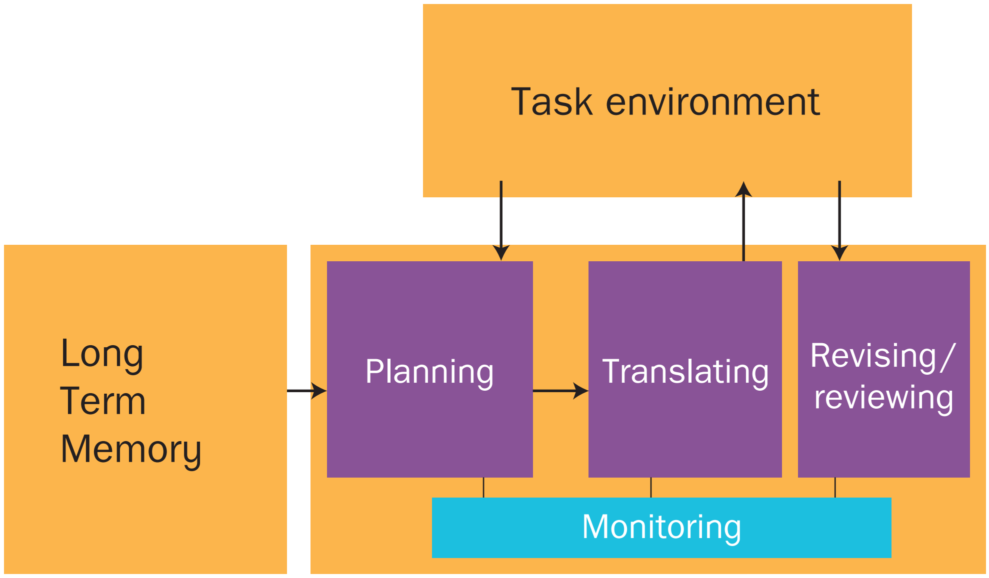

Figure 12.9 The Hayes and Flower model of writing. The model aimed to identify the key cognitive aspects of the writing process.

Source: Adapted from Hayes J. R., & Flower L. S. (1980). *Identifying the organization of writing processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

The second component of the Hayes and Flower model concerns the writer's long-term memory. The availability and accessibility of relevant information in long-term memory supports the writing process. In addition to knowledge about the subject matter, the long-term memory system stores schematic information that will shape the writer's view of readers' expectations. The third component concerns the writing process itself and the immediate cognitive demands it brings. Working memory demands (see Chapter 4) are relevant to this component.

Hayes and Flower discuss three general stages of writing: planning, translating and reviewing (see Figure 12.9). Planning includes the sub-operations of generating, organizing and goal setting. Translating converts ideas from memory into sentences on paper. Reviewing involves reading and editing. Revision is a key stage in the writing process; good writers revise more, and focus on the meaning of the text rather than the more superficial qualities of the writing (Flower et al., 1986). This model considers writing as a metacognitive act, with an executive process monitoring the key processes of planning, translating and revising (Peverly, 2006). The goal of writing is to create 'reader-based prose', as opposed to 'writer-based prose' (Peverly, 2006). Hayes's (1996) revised model of writing acknowledges the central role of working memory capacity in skilled writing.

Evaluation

The Hayes and Flower model brings together three key aspects of writing: the writing task itself, the cognitive processes involved in writing, and the writer's knowledge and long-term memory. The model moved away the previous sequential models and placed an emphasis on multiple cognitive sub-processes that allow writers to plan, revise and re-draft text. Later versions of the model also considered the role of working memory in writing. But a number of questions remain. What aspects of long-term and working memory function predict writing quality? How might memory processes support the developing writer? How does cognition differ in expert and novice writers? What role does oral language fluency play in the writing process? What role do reading skills play in the writing process? The element of time is also absent in this and other cognitive accounts of writing (see Becker, 2006): when do the various cognitive activities occur, when are certain actions initiated?

Models of writing typically have not considered the role of lower level processes in the development of writing skill, yet clearly the mechanical aspects of writing (e.g. motor skills) are an important support to the higher level cognitive processes. Box 12.7 considers how physical writing speed affects cognitive processing.

Box 12.7 Practical Application: Taking lecture notes – speed predicts quality

As a student, your writing skills are tested thoroughly throughout your studies, with examinations, essays and practical reports all demanding the type of high-level planning and execution described in the Hayes and Flowers model above. But what of writing tasks such as lecture note taking? How might aspects of that writing task affect learning?

Peverly (2006) summarizes the key skills involved in taking lecture notes. First, information must be held in verbal working memory and this is subject to capacity restrictions (see Chapter 5). Second, the key points must be selected from the information held in working memory, and third, those key points must be transcribed before they are forgotten; this requires efficient writing. All of this must take place while attention is maintained on the ongoing lecture. Although many individual differences come into play here (including working memory capacity, verbal ability, etc.), Peverly (2006) was most interested in writing speed itself. If you can write the key points down quickly, will later test performance be facilitated?

Peverly and colleagues had students listen to a lecture on problem solving and take notes. Measures of transcription speed, working memory, spelling, and identification of salient information were also taken. Their data showed that faster handwriting speed was associated with higher quality lecture notes. By practising the basic processes that allow us to take notes, we increase note-taking speed, and this frees up working memory for attending to the higher level processes involved in selecting salient information, following the argument being made in the lecture, integrating the current information with previous points made, and so on. As Peverly summarizes, 'the best way to enhance the efficiency of a limited-capacity processing system is through instruction and practice, especially of basic skills, so that the capacity of working memory can be devoted to the higher order skills necessary to achieve academic goals' (2006, p. 209).

These data suggest that interventions designed to improve transcription fluency may lead to improved lecture note taking and thereby improved test performance.

Summary

In this chapter we have considered the nature and components of language and the cognitive processes involved in the production of speech. Language is our principal means of communication and seems to be uniquely human. While language shares properties with other animal communication systems, no animal system has all its features. The special features of language are productivity, displacement and duality of patterning.

The components of language are phonemes, morphemes, syntax and discourse. The basic sound units of a language are phonemes; its meaning units are morphemes. Sentences are composed of morphemes and are structured using syntax.

Speech production involves four main stages. Stage 1, conceptualization, prepares a thought for conversion into language. The second stage involves the formulation of a linguistic plan. Formulation also involves syntactic planning; during this sub-stage the order in which the selected words will be output is decided. The third stage involves the articulation of the plan. During this stage the sounds for the word are accessed and articulated. In a final fourth stage, the output is monitored so that corrections can be made if errors occur. Models of speech production differ in terms of the degree of modularity and interaction said to occur between processing levels.

This chapter examined speech errors and their contribution to our understanding of speech production. Speech errors occur in a number of types and are not random. They support the idea that the production of speech involves a number of distinct stages.

This chapter also examined the language deficits that follow brain injury in adults. The patterns of deficit in Broca's aphasia, Wernicke's aphasia and anomia suggest a dissociation between syntactic/output and semantic/comprehension processes in language processing.

Finally, we considered language production in the form of writing, and the three stages of the writing (as composition) process: planning, translating and reviewing.

Review questions

1. What are the main features of human language?
2. What do the acquired disorders of language contribute to our understanding of normal speech production?
3. How does the analysis of speech errors contribute to our understanding of normal speech production?
4. What are the key differences between modular and interactive accounts of speech production?
5. How do the processes of writing differ from those of speech production?

FURTHER READING

Brown, R., & McNeill, D. (1966). The 'tip of the tongue phenomenon'. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*, 73–111.

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, *130*(5), 1432–1441.

Harley, T. A. (2010). *Talking the talk: Language, psychology and science*. Hove: Psychology Press.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*, 88–96.