Defining listening
Section Introduction:
The nature of processing

This section defines listening in terms of overlapping types of processing: neurological processing, linguistic processing, semantic processing, and pragmatic processing. A complete understanding of listening needs to account for all four types of processing, indicating how these processes integrate and complement each other.

Chapter 1 describes neurological processing as involving consciousness, hearing, and attention. The chapter describes the underlying universal nature of neurological processing and the way it is organised in all humans, for users of all languages. The chapter also attempts to elucidate nature of individual differences in neurological processing, to explain the individualised nature of the listening experience.

Chapter 2 describes linguistic processing, the aspect of listening that requires input from a linguistic source – what most language users would consider the fundamental aspect of listening to language. This chapter begins with a section on perceiving speech, and proceeds to describe the way in which listeners identify units of spoken language, use prosodic features to group units of speech, parse speech into grammatical units and recognise words.

Chapter 3 details semantic processing, the aspect of listening that integrates memory and prior experience into understanding events. This chapter focuses on comprehension as constructing meaning and the memory processes that are involved.

Chapter 4 introduces the broad issue of pragmatic processing. While closely related to semantic processing, pragmatic processing evolves from the notion of relevance – the idea that listeners take an active role in identifying relevant factors in verbal and non-verbal input and inject their own intentions into the process of constructing meaning.

Finally, Chapter 5 describes automatic processing – the simulation of listening by a computer. This chapter outlines the ways that natural
language processing by computers emulates the linguistic, semantic, and pragmatic processing of humans.

Section I lays the groundwork for the discussion of teaching listening and researching listening that will follow in subsequent sections. Though a number of teaching and research considerations will be indicated in Section I, the primary focus of the chapters in this section is on understanding the processes themselves.
Neurological processing

1.1 Hearing

A natural starting point for an exploration of listening in teaching and research is to consider the basic physical and neurological systems and processes that are involved in hearing sound.

Hearing is the primary physiological system that allows for reception and conversion of sound waves. Sound waves are experienced as pressure pulses and can be measured in pascals (Force over an Area: \( p = \frac{F}{A} \)). The normal threshold for human hearing is about 20 micropascals – equivalent to the sound of a mosquito flying about 3 m away from the ear. These converted electrical pulses are transmitted from the outer ear through the inner ear to the auditory cortex of the brain. As with other sensory phenomena, auditory sensations are considered to reach perception only if they are received and processed by a cortical area in the brain. Although we often think of sensory perception as a passive process, the responses of neurons in the auditory cortex of the brain can be strongly modulated by attention (Fritz et al., 2007; Feldman, 2003).

Beyond this conversion process of external stimuli to auditory perceptions, hearing is the sense that is often identified with our affective experience of participating in events. Unlike our other primary senses, hearing offers
unique observational and monitoring capacities that allow us to perceive life’s rhythms and the ‘vitality contours’ of events (Stern, 1999) as well as of the tempo of human interaction in real time and the ‘feel’ of human contact and communication (Murchie, 1999).

In physiological terms, hearing is a neurological circuitry, part of the vestibular system of the brain, which is responsible for spatial orientation (balance) and temporal orientation (timing), as well as interoception, the monitoring of sensate data for our internal bodily systems (Austin, 2006). Hearing also plays an important role in animating the brain, what Sollier (2005) calls cortical recharging of the sensory processing centers in the brain.

Of all our senses, hearing may be said to be the most grounded and most essential to awareness because it occurs in real time, in a temporal continuum. Hearing involves continually grouping incoming sound into pulse-like auditory events that span a period of several seconds (Handel, 2006). Sound perception is about always anticipating what is about to be heard – hearing forward – as well as retrospectively organising what has just been heard – hearing backward – in order to assemble coherent packages of sound.

While hearing provides a basis for listening, it is only a precursor for it. Though the terms hearing and listening are often used interchangeably in everyday talk, there are essential differences between them. While both hearing and listening are initiated through sound perception, the difference between them is essentially a degree of intention. Intention is known to involve several levels, but initially intention is an acknowledgement of a distal source and a willingness to be influenced by this source (Allwood, 2006).

In psychological terms, perception creates knowledge of these distal objects by detecting and differentiating properties in the energy field. In the case of audition, the energy field is the air surrounding the listener. The perceiver detects shifts in intensity, which are minute movements in the air, in the form of sound waves, and differentiates their patterns through a fusion of temporal processing in the left cortex of the brain and spectral processing in the right. The perceiver designates the patterns in the sound waves to various learned categories, which is the first stage of assigning some meaning to the sound (Zatorre et al., 2002; Harnad, 2005; Kaan and Swaab, 2002).

The anatomy of hearing is elegant in its efficiency. The human auditory system consists of the outer ear, the middle ear, the inner ear, and the auditory nerves connecting to the brain stem. Several mutually dependent subsystems complete the system (see Figure 1.1).

The outer ear consists of the pinna, the part of the ear we can see, and the ear canal. The intricate funnelling patterns of the pinna filter and amplify
the incoming sound, in particular the higher frequencies, and allows us the ability to locate the source of the sound.

Sound waves travel down the canal and cause the eardrum to vibrate. These vibrations are passed along through the middle ear, which is a sensitive transformer consisting of three small bones (malleus, incus, and stapes) surrounding a small opening in the skull (the oval window). The major function of the middle ear is to ensure efficient transfer of sounds, which are still in the form of air particles, to the fluids inside the cochlea (the inner ear), where they will be converted to electrical pulses and passed along the auditory nerve to the auditory cortex in the brain for further processing.

*Note* The semicircular canals, which are also part of the inner ear, are used primarily for equilibrium but share the same cranial nerve (the eighth) that the auditory system uses, so hearing and balance are interrelated neurally.

In addition to this transmission function, the middle ear has a vital protective function. The ossicles have tiny muscles that, by contracting reflexively, can reduce the level of sound reaching the inner ear. This reflex...
action occurs when we are presented with sudden loud sounds such as the thud of a dropped book or the wail of a police siren. This contraction protects the delicate hearing mechanism from damage in the event that the loudness persists. Interestingly, the same reflex action also occurs automatically when we begin to speak. In this way the ossicles reflex protects us from receiving too much feedback from our own speech and thus becoming distracted by it.

The cochlea is the focal structure of the ear in auditory perception. The cochlea is a small bony structure, about the size of an adult thumbnail, that is narrow at one end and wide at the other. The cochlea is filled with fluid, and its operation is fundamentally a kind of fluid mechanics. (Bioelectric engineers at MIT recently redesigned an ultra-broadband radio chip modelled on the fluid mechanics of the cochlea. See Trafton, 2009.)

The membranes inside in the cochlea respond mechanically to movements of the fluid, a process called **sinusoidal stimulation**. Lower frequency sounds stimulate primarily the narrower end of the membrane, and higher frequencies stimulate only the broader end. Each different sound that passes through the cochlea produces varying patterns of movement in the fluid and the membrane.

At the side of the cochlea, nearest the brain stem, are thousands of tiny hair cells, with ends both inside and outside the cochlea. The outer hair cells are connected to the auditory nerve fibres, which lead to the auditory cortex of the brain. These hair cells respond to minute movements of the fluid in the membrane and **transduce** the mechanical movements of the fluid into nerve activity.

As with other neural networks in the human body, our auditory nerves have evolved to a high degree of specialisation. There are five different types of auditory nerve cells. Each auditory nerve system has different **Characteristic Frequencies (CF)** that they respond to continuously throughout the stimulus presentation. Fibres with high CFs are found in the periphery of the nerve bundle, and there is an orderly decrease in CF toward the centre of the nerve bundle. This **tonotopic organisation** preserves the frequency spectrum from the cochlea, which is necessary for speedy, accurate processing of the incoming signal pulses. Responding to their specialised frequencies, these nerves actually create tuning curves that correspond to the actual shape of their cell and pass along very precise information about sound frequency to the **superior olivary complex** of the central auditory nervous system (Musiek et al., 2007).

The distribution of the neural activity is called the **excitation pattern**. This excitation pattern is the fundamental output of the hearing mechanism. For instance, if you hear a specific sequence of sounds, there is a specific excitation pattern produced in response that is, in principle, precisely the same as the excitation pattern produced in all other hearing humans. While the excitation patterns may be identical, how the hearer
interprets the signal and subsequently responds to it is, of course, subject to a wide range of individual differences, especially age and language learning background.

Concept 1.1  **Excitation patterns and hearing**

Excitation patterns in the inner ear and auditory nerve become automated through experience with familiar stimuli. Without excitation patterns, hearing cannot take place: the auditory stimulus will not reach the brain.

In a sense, this means that not everyone hears the same thing, even though the excitation pattern for a particular stimulus will be neurologically similar in all of us. On a physical level, the difference in our perception is due to the fact that the individual neurones that make up the nerve fibres are interactive – they are affected by the action of all the other neurones they interact with. Sometimes, the activity of one neurone is suppressed or amplified by the introduction of a second tone. In addition, since these nerves are physical structures, they are affected by our general health and level of arousal or fatigue. Another fact that interferes with consistent and reliable hearing is that these nerves sometimes fire involuntary even when no hearing stimulus is present. This occurs when the vestibular nerve, which is intertwined with the auditory nerve and helps us keep our balance, is activated. (Musiek et al., 2007; Moore, 2004).

The physiological properties of listening begin when the auditory cortex is stimulated. The primary auditory cortex is a small area located in the temporal lobe of the brain. It lies in the back half of the Superior Temporal Gyrus (STG) and also enters into the transverse temporal gyri (also called Heschl's gyri). This is the first brain structure to process incoming auditory information. Anatomically, the transverse temporal gyri are different from all other temporal lobe gyri in that they run mediolaterally (towards the centre of the brain) rather than dorsiventrally (front to back).

As soon as information reaches the auditory cortex, it is relayed to several other neural centres in the brain, including Wernicke's area, which is responsible for speech recognition, and lexical and syntactic comprehension, and Broca's area, which is involved in calculation and responses to language-related tasks.

Imaging studies have shown that many other brain areas are involved in language comprehension as well (see Figure 1.2). This neurological finding is consistent with language processing research indicating simultaneous parallel processing of different types of information.

These studies have shown that all of these areas are involved in competent language comprehension to varying degrees, with certain areas more
active while processing particularly complex sentences or disambiguating particularly vague references. Impairments in any one area, often defined as an *aphasia* (if acquired by way of an injury or aging process), can result in difficulties with lexical comprehension, syntactic processing, global processing of meaning and formulation of an appropriate response (Poeppel *et al.*, 2008; Harpaz *et al.*, 2009).

Figure 1.2 *Primary areas of the brain involved in listening.* Several areas of the brain are involved in listening, most of them in the left hemisphere. 

(a) The left prefrontal cortex is involved in processing information during speech comprehension. 
(b) The left pars triangularis is involved in syntactic processing. 
(c) The left pars orbitalis is involved in semantic processing of lexical items; the right pars orbitalis (in the right hemisphere of the brain) is involved in semantic processing of discourse. 
(d) The left superior temporal sulcus (STS) is involved in phonetic processing of sounds; the right STS is involved in processing prosody. 
(e) The left plenum temporale is involved in speech–motor interface. 
(f) The primary auditory cortex is involved in speech perception. 
(g) The secondary auditory cortex (which wraps around the primary auditory cortex) is involved in the processing of intonation and rhythm. 
(h) The left superior temporal gyrus (STG) is involved in semantic processing of lexical items; the right STG is involved in semantic processing at the discourse level.
1.2  Consciousness

Concept 1.2  Consciousness and listening

Consciousness is the aspect of mind that has a self-centred point of view and orientation to the environment. Consciousness is directly related to intentionality – the intention to understand and to be understood.

Once we understand the basic physiology of hearing and listening, we realise that a complex neural architecture underlies our ability to understand language and the worlds around and within us. At the same time, through simple reflection, we realise there are non-physical aspects of processing and understanding that go well beyond the systems we have just outlined.

The concept that has been used most often to describe this neurological-cognitive bridge between individual and universal perception and experience is consciousness (Chafe, 2000). Consciousness is the root concept for describing the processes that initiate attention, meaning construction, memory and learning.

Just as we characterised sound perception as a neurophysical process originating from an energy pattern in air outside of us, we may think of consciousness in a similar way. Consciousness has been described as a flow of energy, emerging when two cognitive processes coincide: (1) The brain identifies an outside object or event as consisting of independent properties; and (2) The brain sets up the listener as the central agent who willingly and purposefully witnesses this object or event. Consciousness is the phenomenon of experiencing this integration as a subjective phenomenon (cf. Czikszentmihalyi, 1992; Chella and Manzotti, 2007).

Beyond this characterisation of subjective experience, it has been said that consciousness is a dynamic neurophysiological mechanism that allows a person to become active and goal-directed in both internal and external environments (Alexandrov and Sams, 2005). This means that consciousness is a continuous force that links experiences in the internal and external environments and allows the experiencer to make sense of these experiences and, to some degree, direct them.

For the purposes of describing listening, the concept of consciousness is important because it helps to define the notion of context. Consciousness involves the activation of portions of the listener’s model of the surrounding world – a model that is necessarily self-referenced. The portions of this model that are activated are those that are involved in understanding the current encounter, including whatever language is associated with it. Viewed technically, this active portion of the model is constructed from
perceptual contact with the external event (external context) and from our subjective experience (internal context).

The concept of consciousness is important for communication – both listening and speaking – because something must direct the individual’s attention to the external world. For the speaker, consciousness influences what aspects of the person’s experience to communicate – the signalling and displaying levels of communication (Holmqvist and Holsanova, 2007). For the listener, consciousness guides the person’s intentions to experience the speaker's world and to attempt to construct meaning from this experience.

**Concept 1.3  The properties of consciousness**

There are five properties of consciousness that affect listening.

- Consciousness is *embedded* in a surrounding area of peripheral awareness. The active focus is surrounded by a periphery of semi-active information that provides a context for it.
- Consciousness is *dynamic*. The focus of consciousness moves constantly from one focus, or item of information, to the next. This movement is experienced by the listener as a continuous event, rather than as a discrete series of ‘snapshots’.
- Consciousness has a *point of view*. One’s model of the world is necessarily centred on a self. The location and needs of that self establish a point of view, which is a constant ingredient of consciousness and a guide for the selection of subsequent movements.
- Consciousness has a need for *orientation*. Peripheral awareness must include information regarding a person’s location in space, time, society and ongoing activity. This orientation allows consciousness to shift from an immediate mode, in which the person is attending to present, tangible references, to a distal mode, in which the person is attending to non-present, abstract, or imaginary references and concepts.
- Consciousness can *focus* on only one thing at a time. The limited capacity of consciousness is reflected as a linguistic constraint: A speaker can produce only one focus of consciousness at a time, which is reflected in brief spurts of language, called intonation units.


**1.3 Attention**

Attention is the operational aspect of consciousness and can be discussed more concretely. Attention has identifiable physical correlates: specific areas of the brain that are activated in response to a decision to attend
to a particular source or aspect of input. Attention is the focusing of consciousness on an object or train of thought, which activates parts of the cortex that are equipped to process it (Figure 1.3).

Because of the deliberate nature of attention, we can consider attention to be the beginning of **involvement**, which is the essential differentiation

![Diagram of stages of attention](image)

**Figure 1.3 Three stages of attention.** Attention consists of three nearly simultaneous stages.

- **Stage 1 is arousal:** in response to a stimulus (internal or external) neurotransmitters originating in the brain stem (reticular activating system) fire throughout the brain, activating brain chemicals (dopamine and noradrenaline) and creating bursts of electrical activity.

- **Stage 2 is orientation:** the superior colliculus regulates the neurotransmitters and directs them to areas of the brain that will be used for processing the stimulus.

- **Stage 3 is focus:** the lateral pulvinar region of the brain (the part of the brain most active in experiences of consciousness) locks the neurotransmitters onto the parts of the cerebral cortex needed to process the stimulus.
between hearing and listening. Psychologists often refer back to the original definition given by William James, considered the founder of modern experimental psychology.

**Quote 1.1  William James on attention**

Everyone knows what attention is. It is the taking possession of the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. *Focalisation and concentration of consciousness* are of its essences. It implies withdrawal from some things in order to deal effectively with others.

William James (1890: 405)

In neurolinguistic research, attention is seen as a timed process requiring three neurological elements: **arousal, orientation and focus**. Arousal begins with the **Reticular Activating System (RAS)** in the brain stem becoming activated. When this happens, the RAS releases a flood of neurotransmitters to fire neurons throughout the brain. Orientation is a neural organisation process performed near the brain stem (specifically, in the superior colliculus part of the brain above the brain stem). This process engages the brain pathways that are most likely to be involved in understanding and responding to the perceived object (i.e. the external event or the internal train of thought). Activation is simultaneous on both sides of the brain – in the right hemisphere, which functions as a parallel processor, and in the left hemisphere, which functions as a serial processor. Focus is achieved in the higher cortex of the brain, the lateral pulvinar section. This process selectively locks on to the pathways that lead to the frontal lobe of the brain and are involved in processing incoming stimulus, thus allowing for more efficient use of energy (Carter, 2003).

Two notions are central to understanding how attention influences listening: **limited capacity** and **selective attention**. The notion of limited capacity is important in listening. Our consciousness can interact with only one source of information at a time, although we can readily and rapidly switch back and forth between different sources, and even bundle disparate sources into a single focus of attention. Whenever multiple sources, or streams, of information are present, **selective attention** must be used. Selective attention involves a decision, a commitment of our limited capacity process to one stream of information or one bundled set of features.
As we listen, our attention can be selectively directed to a rich variety of acoustic features beyond linguistic aspects, including spatial location, auditory pitch, frequency or intensity, tone duration, timbre, and characteristics of individual voices. Depending on which of the multiplicity of acoustic dimensions we choose to attend to, a different area of the brain will become active. Indeed, it has been shown that the locations of the multiple loci of attentional influence on auditory information processing are flexible and dependent not only on the nature of the input, but also on the specific demands of the behavioural task being performed. Another influence on the cortical locus of attention is the involvement of other modalities. For instance, if visual and auditory attention are activated simultaneously, different areas of the frontal-parietal network in the brain will become involved.

Concept 1.4  Processes of attention

- Attention is a limited capacity system.
- Automatic activities that require little or no attention do not interfere with each other.
- Controlled processes require attention and interfere with other control processes.
- Attention can be viewed as three separate but interrelated networks: alertness, orientation and detection.
  1. Alertness represents a general readiness to deal with incoming stimuli.
  2. Orientation refers to a specific aligning of attention.
  3. Detection is the cognitive registration of sensory stimuli.
- Detected information is available for other cognitive processing.

Concept 1.5  Selective attention and processing breaks

Among the best known experimental studies dealing with selective attention are dichotic listening studies in which subjects are presented with different messages through left and right earphones. When told to attend to one message only or shadow it, subjects can readily comply, switching attention to the second message. However, subjects can shift attention only at pauses in the attended message, which suggests that we can shift our attention only at suitable ‘processing breaks' in the input.

Just as important, results from these studies show that attention is needed not only for monitoring input, but also for effectively storing and retrieving messages. A consistent finding in these experiments is that only information in the attended channel (i.e. the ear with the attended input) can be remembered.
An everyday example of this is the **cocktail party effect**. In a chaotic – inherently unpredictable – cocktail party environment, numerous streams of conversation are taking place, yet you can attend to only one at a time. It is possible to focus on a conversation taking place across the room while ignoring a conversation that is closer and louder. Attention is directional and under the control of the listener, within certain constraints. This ability is also much reduced in individuals with hearing loss, or with hearing aids and cochlear implants.

Although attention can usually be controlled, shifts in attention are not always voluntary. For example, while we are watching television, our baby starting to cry takes over the attention system momentarily whether we want it to or not. Instinctively, we respond to what is perceived to be most relevant to our needs. Beyond obvious examples of overt emergency signals (such as a baby’s crying signalling a need for us to take care of it) overtaking our previous attentional focus (such as watching the news), our needs are complex and subtle and may be ordered in ways that are not fully conscious to us. Because of this complex nature of our informational and emotional needs, we may often respond to subtle distractions when we are listening and become derailed from our original focus.

### 1.4 Individual differences in neurological processes

Among linguists, psychologists, and philosophers, language is regarded as the most complex of all human behaviours. And within the modalities of language use, speech processing may be the most intricate. At any given moment during language processing, we may be engaged simultaneously in speaking, hearing, reading, formulation and comprehension. Each of these individual component skills requires the involvement of large areas of the brain and a complex interplay of neural health, attentional readiness, local neural processing, coordination of functional neural circuits, and high-level strategic organisation. As we have seen in earlier sections of this chapter, work in cognitive neuropsychology has helped identify the basic functions of brain areas in terms of language processing. New scanning techniques also are leading to a fuller understanding of these interactions and how they are linked together into functional neural circuits for language processing.

In spite of these common capacities for language processing, not all humans process language in the same way. As in other areas of neural processing, individuals display a great range of differences across these functions. This section outlines six critical differences among individuals:

- **Local processing.** In terms of basic-level processing, individuals show marked differences in basic attributes such as speed of neural transmission,
activation of neural transmitters, involvement of the thalamus (relay centre for all neural impulses) and hippocampus (part of the limbic system involved in orientation), memory and attention, and patterns of neural connectivity.

- **Commitment and plasticity.** As basic linguistic functions develop, they become confined to progressively smaller areas of neural tissue, a process called **neural commitment**. This leads to a beneficial increase in automaticity and speed of processing, but it also results inevitably in a decline in plasticity. (There is also some loss in the potential to function if brain injury occurs in an adjacent area). It appears that the process of neural commitment leads to a neural separation between different languages in bilinguals and second language learners. The plasticity or neural flexibility required for language reorganisation declines progressively through childhood and adolescence and may be the primary cause of some of the difficulties that adults face in second language learning (Gitterman and Datta, 2007; Van Den Noort et al., 2010).

- **Integrative circuits.** Current models of the formation and consolidation of episodic memories focus on the role played by the hippocampus in forming integrated representations (MacWhinney, 2005a; Kroll and Tokowitz, 2005). In terms of language learning and use, these neural connections allow a variety of local areas of the brain to form a series of impressions of sensory and conceptual aspects of an utterance, which are then linked into a new grammatical form or syntactic construction. (All mammals use connections between the hippocampus and local areas to form memories. However, humans are unique in using those connections to support language learning.) In addition to this central memory consolidation circuit, a variety of local circuits are likely used in analysing and breaking apart local memories through a process called **resonance** (Grossberg, 2003). Resonant circuits copy successfully detected linguistic forms to temporary local buffers so that the system can focus on incoming, unprocessed material while still retaining the recognised material in local memory. As with all neural mechanisms, differences in the efficiency of these individual circuits can be assumed.

- **Functional neural circuits.** The types of local integration supported by the episodic memory system are complemented by a variety of other functional neural circuits that integrate across wider areas of the brain. A prime example of such a circuit is the **phonological rehearsal loop** (Lopez et al., 2009), which links the auditory processing in the temporal lobe with motor processing from the prefrontal cortex. We use this loop to store and repeat a series of words or to speed the learning of new words. Differences in the abilities of listeners to store items in this loop have been shown to correlate strongly with relative success in both L1 and L2 learning (Aboitiz et al., 2010; Gathercole et al., 1994).
• **Strategic control.** Brain functioning can be readily modified, amplified, integrated and controlled by higher-level strategic processes. These higher-level processes include mood control, attentional control, motivational control as well as learning strategies and applications of cognitive maps and scripts. The degree to which the listener can activate and apply these higher-level processes will determine relative success and failure in language comprehension in specific instances and in long-term acquisition (Van Heuven and Dijkstra, 2010).

• **Level of attention.** Some listeners pay more attention to overall conceptual structure, attempting to process incoming language more through top-down inferential, whereas other learners focus more on bottom-up detail (Bransford, 2003). This individual difference is also likely to be important in determining the relative success of listeners in language comprehension to specific texts and in longer-term acquisition of the language.

**Summary: organisation of neurological processing**

This chapter has surveyed the neurological processes that are involved in listening. Though the processes are wired through complex electrochemical circuitry, these processes are far from mechanist and robotic. We humans are a meaning-oriented species, and our neurobiology is geared not only to process information and make sense of the external world, but also to understand and find meaning in both the external world and our internal world.

Philosophers have long argued that the deepest sources of human understanding lie not in external information sources or information processing, but in feelings, emotions, qualities and patterns of bodily perception and motion. Images, qualities, emotions and metaphors are rooted in our physical encounters with the world and provide the basis for our most profound feats of abstract understanding. As Johnson (2007) emphasises, though the contemporary study of neurolinguistics often focuses on the more scientific aspects of information processing and meaning building, we should not lose touch with the understanding that meaning-making is also fundamentally human, interactive, and aesthetic.