Psycholinguistics: Overview

A Anderson, University of Glasgow, Glasgow, UK

Emergence of Psycholinguistics in the Late 1950s and 1960s from the Chomskyan Revolution

Although the study of language has been part of psychology from its earliest years, including for example in the work of Wilhelm Wundt, the father of psychology, a distinct field of psycholinguistics emerged in the late 1950s largely in response to the impact of Chomsky. In the preceding decades, notably in the United States, psychology had been dominated by the behaviorist approach of researchers such as B. F. Skinner. They treated language as a form of verbal behavior, which, like all other behavior, they believed was governed by simple stimulus–response associations. Chomsky demonstrated the shortcomings of the behaviorist approach in explaining the productivity of language and its complexity, and his work, notably *Syntactic structures* (1957), provided a major impetus for a new kind of psychological investigation of language. This was driven by an interest in the mental representation of language in general and syntactic structures in particular (*see* **Psycholinguistics: History**).

Psychology since the demise of behaviorism has again been concerned with understanding the way that people accomplish various information-processing tasks. In the field of psycholinguistics this means a concern with the cognitive processes by which a string of sounds in an utterance, or marks on a page, are processed to identify individual words and sentences, and how this emerging structure becomes mentally represented as a meaningful concept. The goal of this process is to derive models that account for how people achieve this so rapidly and successfully, given what we know about the general limitations of human cognitive processing. To oversimplify: the psychologist is concerned with **how** the linguistic units are processed and represented; the linguist is concerned with the description of the structures that emerge from any such processes.

Research Topics in the Early Years of Psycholinguistics

In its early years psycholinguistics reflected the concerns of linguistics and the central role of syntax. Psychologists such as Miller and Isard (1963) showed that the syntax influences the way people interpret sentences, and even how many words people can remember from a string of words that make no sense. More words are remembered from a 'sentence' like 'Accidents carry honey between the house' than from strings with no syntactic structure such as 'On trains hive elephants the simplify.' The focus on syntax and its importance in language processing led many psycholinguists to try to test the psychological reality of Chomsky's theory of generative grammar. Experiments were designed to explore the notion that when people process sentences, what they are doing is retrieving the deep syntactic structure as described in Chomsky's transformational grammar. So initially, a number of studies seemed to show that the relative ease or difficulty with which a reader or listener could process a given sentence was directly related to its syntactic complexity. Chomsky's kernel sentences, equivalent to active affirmative declarative sentences, were recalled most easily and processed most quickly, while sentences including one or more transformations such as negative or passive forms were more difficult to process. These kinds of study led to the so-called derivational theory of complexity. This was superseded as it became clear that the results of many of the studies that apparently provided support for this purely syntactic view of how people process sentences could also be accounted for by the influence of semantic factors. Although sentence processing remains one of the most significant topics in psycholinguistics, the range of language phenomena that are studied has broadened considerably since the early 1960s. The assumption that the role of psycholinguistics is to demonstrate the psychological validity of any particular syntactic theory has also been overtaken.

Models of Sentence Comprehension

Many of the models of sentence comprehension that have been developed in psycholinguistics try to elucidate the cognitive processes that are involved when a reader or listener interprets a sentence. There is considerable experimental evidence that sentence comprehension is incremental, that is, that an interpretation is built up on a moment-by-moment basis from the incoming linguistic information. The evidence for this incremental processing is particularly striking in the way listeners recognize spoken words (see below), which are often identified before all the acoustic information has been heard.

Even in written language processing we have clear signs of the incremental nature of linguistic processing. This is illustrated by the difficulties most readers have with sentences like the following: 'The horse raced past the barn fell' (Bever, 1970). This is known as a garden path sentence, because nearly all readers interpret this **while** they are reading, with 'raced' as an active verb and so expect the sentence to end after 'barn.' They do not realize that the sentence could have an equivalent interpretation to 'The horse that raced past the barn, fell.' The incremental way that sentences seem to be interpreted by readers or listeners is a major source of potential ambiguities of interpretation. Many sentences in a language have potentially more than one interpretation as they are processed, yet the reader or listener is usually not aware of any problem in arriving at one clear interpretation of a sentence.

The cognitive architecture that underpins such an achievement has been a source of much debate in psycholinguistics. Some researchers have held that in the frequent cases where more than one analysis of a sentence is possible, the reader or listener computes all possible analyses in parallel. The difficulty of garden path sentences has led others to propose serial models, where it is assumed that a single analysis is computed and corrected later if this is needed. Models have been proposed to account for the empirical evidence on sentence-processing difficulties. These often involve various versions of parallel analyses, where candidate analyses are only active for a brief period of time or are ranked according to frequency in the language or plausibility with the context. One of the most influential of these accounts is the constraint-based model of MacDonald *et al.* (1994). The weightings attached to each candidate analysis are based on the frequency of a syntactic structure in the language, the plausibility of the words in the sentence to their assigned syntactic roles, etc. So this model is an example of an interactive model where many different sorts of information, syntactic, semantic, pragmatic, contextual, and frequency, can all play simultaneous roles by activating alternative interpretations of the incoming linguistic information. In models of this type, semantic factors can override syntactic processing biases. This model is attractive to psychologists for a number of reasons. It is amenable to modeling by connectionist approaches (*see* **Cognitive Science: Overview**) and it avoids the problem of having to base cognitive processes on syntactic rules, which for psychologists appear to change arbitrarily with changes in linguistic theories.

In contrast, one of the other most influential models of sentence processing is the garden path model (Frazier, 1979). This uses only syntactic principles in its initial stage. An analysis is computed based on two syntactic preferences, the most important being the principle of minimal attachment, the other being the principle of late closure. The first principle means that the parsing of the sentence that produces the simplest parse tree, with fewest nodes, takes precedence. So a sentence like 'Mary watched the man with the binoculars' is usually interpreted to mean that Mary (not the man) was using binoculars. According to Frazier's interpretation of phrase structure rules, this interpretation involves one node fewer than the alternative and so demonstrates the principle of minimal attachment in action. This principle takes precedence over the principle of late closure. This is the preference to attach incoming materials to the current phrase or clause. This latter principle is used to explain the preference for interpreting sentences like 'John said he will leave this morning' to mean that the phrase 'this morning' relates to the verb 'leave,' not the verb 'said.'

As part of the ongoing debate about the adequacy of different models of the parsing process there has been an active discussion in the experimental literature over several years about the extent to which semantic factors can override or guide the analysis of syntactic structure. This has been explored in several studies focusing on the ease or difficulty with which sentences containing reduced relative clauses can be processed. Several studies have tested how people interpret sets of sentences like the following:

(1a) The defendant examined by the lawyer turned out to be unreliable.
(1b) The defendant that was examined by the lawyer turned out to be unreliable.

(2a) The evidence examined by the lawyer turned out to be unreliable.
(2b) The evidence that was examined by the lawyer turned out to be unreliable.

Sentences like (1a), which contain a reduced relative clause, are more difficult to process than their equivalent full relative clause (1b). Readers initially treat 'examined' as a main verb whose subject is 'the defendant.' They then have to reanalyze this garden path when they reach the phrase 'by the lawyer.' The argument is the extent to which structurally similar sentences (e.g., [2a]) cause readers to have equivalent processing problems. This is what might be expected from a purely syntactic view of parsing. In contrast, in an interactive constraint-based model, the semantic implausibility of interpreting an inanimate noun such as 'evidence' as the subject of the verb 'examined' should protect the reader from the need to reanalyze an initial incorrect syntactic structure.

Trueswell *et al.* (1994) seemed to show just such a pattern. This was considered powerful evidence in support of interactive constraint-based models of sentence processing. More recently, Clifton *et al.* (2003) have challenged the evidence that semantic factors override syntactic processing in the initial stages of parsing. They used more sophisticated techniques for monitoring and analyzing eye movements to determine the processing difficulties experienced by readers. They found that reduced relative clauses caused disruption to processing, irrespective of the semantic plausibility of the relationship between the apparent subject and main verb. Semantic factors, however, influenced how quickly the readers **recovered** from their wrong analysis of the syntax of the sentence. Clifton *et al.* (2003) stressed that the key thing for psycholinguistic models of sentence processing is not whether the data on processing reduced relative clauses support garden path or constraint-based models. They claim that the important goal is to develop parsing models that deal both with the task of creating structure and evaluating the structure that is created.

Although there have been numerous studies of sentence processing conducted by psycholinguists over the decades, the vast majority of these have focused on how readers interpret written sentences. A few studies have tackled the issue of how listeners use the cues in spoken language during parsing. Minimal attachment strategy can be shown to be overcome by the prosodic cues in real spoken sentences. Similarly, the principle of late closure, which can cause syntactic ambiguities in written sentences, can be less problematical in spoken materials because of clear prosodic cues to the intended interpretation. Even the apparent errors in spoken sentences, such as the disfluency 'uh,' can have an impact during the parsing process. Bailey and Ferreira (2003) presented sets of sentences to listeners such as the following:

(3a) Sandra bumped into the busboy and the uh uh waiter told her to be careful.
(3b) Sandra bumped into the busboy and the waiter uh uh told her to be careful.

These sentences are ambiguous up to the point when the listeners hear 'told.' 'Sandra' could have bumped into 'the busboy' or 'the busboy and the waiter.' Yet the listeners who heard the materials most often interpreted sentences like (3a) to mean that 'the waiter' was the subject of a new clause, i.e., that it was the subject of the verb 'told.' This shows that disfluencies can systematically influence the way listeners parse incoming sentences. The same effect was observed when the interruption was not a disfluency but an environmental sound such as a telephone ringing.

Speech Production and Speech Errors

These studies represent a welcome aspect of the broadening of the psycholinguistic research agenda to include more consideration to the production and comprehension of spoken as well as written language. The study of speech disfluencies is one part of this. Speech disfluencies encompass a range of phenomena, including pauses in speech such as silences, filled pauses, and fillers such as 'uh' and 'um,' as well as speech errors such as slips of the tongue, spoonerisms, and malapropisms. When we speak we aim to produce a grammatically well-formed utterance with no noticeable hesitations. Yet this ideal delivery cannot always be achieved. It is estimated that around 5% of words in speech are disfluent in some way. Yet these disfluencies are not random in their patterns of occurrence. Even the similar-sounding disfluencies 'uh' and 'um' have been shown to have systematic contexts of use. Speakers use 'uh' before a short delay in their speech production but use 'um' before a more significant delay. Speakers seem to become disfluent because they are experiencing some kind of problem in planning and producing their utterance. Speakers have been found to pause more before unpredictable words, suggesting they might be experiencing word-finding difficulties. Speakers also pause more at the start of an intonation unit, which suggests that pausing is related to the speech-planning process (*see* **Pauses and Hesitations: Psycholinguistic Approach**).

Speech errors have also been studied by psycholinguists, who have classified them according to the assumed units of processing and types of mechanism involved in their production. For example, speech errors can relate to the phonemic features of the word, the syllabic structure, or the phrase or sentence being produced. This is illustrated in one of the famous errors reportedly produced by Dr Spooner in the 19th century when rebuking one of his students: 'you have tasted the whole worm' when he presumably intended to say 'you have wasted the whole term.' Here the initial phonemes are swapped but the rest of the morphemic and syntactic structure of the target utterance is preserved. Errors of this type became known as 'spoonerisms' as a result.

The kinds of error that occur can tell us a good deal about how speech is produced. Speech errors are very varied. They can reflect many different linguistic levels. Errors can involve the sounds of the words involved, for example saying 'the lust list' for 'the lush list.' They can relate to the intended words in a phrase, 'the pin of a head' being said in place of 'the head of a pin.' Errors may also focus on the semantic relations of the intended words, so a speaker may produce 'I like berries with my fruit' rather than 'I like berries with my cereal,' among many other forms of errors.

Some types of errors, however, do not occur and these patterns of occurrence and nonoccurrence have been used to help understand the speech production process. Content and function words are not substituted for one another and indeed substitute words are usually the same part of speech as the intended target word. When the wrong sound is produced, however, the substituted phoneme seems to have no grammatical relation with the intended target sound. The spacing between the errors is also informative for models of production. Errors of word substitution usually involve words that are a phrase apart. Yet sound errors seem to relate neighboring words. This sort of evidence has been used by several researchers to develop general models of the speech production process (*see* **Speech Errors: Psycholinguistic Approach** for more details).

Speech Recognition

Psycholinguists have also been concerned with exploring the processes involved in speech perception. Jusczyk and Luce (2002) summarized over 50 years of research on this topic. They described key research issues in the domain as understanding invariance, constancy, and perceptual units. In speech there are no invariant acoustic features that map directly to corresponding phonetic segments. The acoustic properties of sounds vary widely depending on the surrounding linguistic context. To make matters even more complex for the listener, there is also wide variability in the way phonetic segments are produced by different speakers depending on age, sex, and individual speaker characteristics. It is not even easy to determine what are the basic perceptual units that listeners use to recognize speech. Some research studies seem to suggest the phoneme as the basic building block of perception while others show advantages of the syllable over the phoneme. Conversational speech is therefore a very variable signal that does not even provide clear cues to boundaries between words. Yet understanding words in speech is an effortless and successful process for listeners with normal hearing. How is it done?

One of the key research challenges for psycholinguists was therefore to produce models of spoken word recognition. One of the most influential models is the Cohort Model (Marslen-Wilson and Welsh, 1978; Marslen-Wilson, 1989). In this account, when a listener hears an initial sound, all the words known to start with that sound become activated. This cohort of candidate words is gradually whittled down to a single word, as more acoustic information is processed and candidate words are eliminated. An important feature of the model is the uniqueness point. This occurs when the listener has heard enough acoustic information to reduce the cohort to a single candidate, i.e., there are no other words known to the listener with that particular sequence of phonemes. Syntactic and semantic context from the surrounding discourse can also play a role in rejecting potential candidate words. In later versions of the model, this can only occur after the uniqueness point, though in earlier versions, context could also be used to eliminate possible words before the uniqueness point. The recognition point occurs when a single item remains in the cohort. This may be before the end of a word.

One of the strengths of the model is the way it can account for the speed with which spoken word recognition often occurs, often occurring before the word offset. In later versions of the model the activation of candidate words is a graded process. Candidate words are not completely eliminated as acoustic information accumulates, but rather they have their activation level reduced. This can rise again if later acoustic information matches a rejected candidate word. This is important, as one of the problems of conversational speech is that individual words or phonemes are often not articulated clearly and the listener must be able to recover and recognize words whose initial sounds were mispronounced or misheard in a noisy environment.

The main competitors to speech recognition models are various connectionist accounts, notably Trace (McLelland and Elman, 1986). This is a connectionist or parallel distributed processing (PDP) account of spoken language processing. In such models researchers were attempting to produce computational models of language processes which were inspired by what was known about the structure and processes of the human brain (*see* **Cognitive Science: Overview**). In spoken word recognition models, this meant that the feature units were simple units with dense series of interconnections, which processed by sending many messages in parallel. Such structures were developed as analogous to the neural architecture of the brain with its many nerves and interconnections.

The Trace model has three layers of units, corresponding at the lowest level to features, then at the next layer to individual phonemes, then at the top layer to complete words. All have dense arrays of interconnections between them, which like nerve pathways can be excitatory or inhibitory. The connections between levels are excitatory and connections within levels are inhibitory. Connections between levels operate in both directions, so both top-down and bottom-up processing can occur. The connectivity of Trace means that evidence is boosted and plausible hypotheses about the possible words emerge strongly. So a feature such as voicing will energize the voice feature units. These will then transmit activation to all the voiced phonemes at the phoneme level, which will in turn activate all the words that begin with these phonemes. At each level the activated units inhibit competitors at the same level, so reducing possible competitors. A word is recognized when in the end a single active unit remains.

Trace is a very interactive model. It gives context a bigger role in recognition than the cohort model. As a computational model it has the virtues of specificity. Trace models can be built and simulations run and then compared with the experimental data from human listeners. These comparisons have generally shown that Trace can cope with some of the problems of variability in production of features and phonemes, can account for the context effects on spoken word recognition that have been reported in the literature, and can cope with the kind of degraded acoustic input that is so typical of real conversational speech. The main criticisms that have been leveled at Trace concern the large role that context is given, which may be an overstatement of how this operates in human recognition. The other limitation is the less than elegant way that the time course of speech recognition is modeled, with duplication of the levels and nodes over successive time periods.

Other connectionist models have also been developed. Despite the competing architectures of the models, there is general agreement on key aspects of the speech recognition process. This involves activation of multiple candidate words, followed by competition among those known lexical items that share a similar sound profile; these processes have to be able to cope with less than ideal acoustic input and deliver perceptions of words very rapidly (for further details *see* **Speech Recognition: Psychology Approaches**).

Discourse Processing

The gradual broadening of the psycholinguistic research agenda has not been limited to the inclusion of speech alongside written language. Psycholinguists have also shown a growing interest in the processes involved in the interpretation not just of single sentences, but of complete texts. One of the first psychologists to explore the complexities of this process was Bartlett (1932). In his research on the way people remembered and reproduced stories that they had heard, he highlighted key research themes in discourse processing that are still current research topics. Bartlett noted how quickly the surface form of the story is lost and an individual's own interpretation of the text is what is remembered. He introduced the concept of a 'schema,' which was used when readers recalled a narrative. This consisted of an organized set of information based on prior experiences that is used in interpretation and recall. The interpretation of a narrative that is retained consists of a mix of input from the text and from schemata. In some of Bartlett's studies, British students listened to North American folktales. When asked to recall the stories accurately, strange narrative details relating to the activities of ghosts were unconsciously altered and supplemented by details from the participants' world knowledge, so that the recalled version of the stories became more coherent by conventional Western standards.

More recently, a growing body of psycholinguistic research has been addressing the challenges of how readers build up a coherent mental representation to create a sense of the narrative world. One of the key concerns of psycholinguistic studies of text or discourse processing is the problem of inferences. Since Bartlett's seminal studies, it has been known that readers expand on what is in the text by drawing inferences based on their knowledge of the world. But what are the time course and limits of such inferencing? Many studies have been concerned with addressing such questions. Experimental studies have shown that some inferences seem to be made automatically **as** we read a text while some are made **later** to resolve apparent problems or inconsistencies. This was demonstrated in a study by Sanford and Garrod (1981), who presented readers with pairs of sentences such as the following:

(4a) Mary was dressing the baby.
(4b) The clothes were made of wool.

(5a) Mary was putting the clothes on the baby.
(5b) The clothes were made of wool.

The participants read the second sentence just as quickly in both cases. This suggests that verbs such as 'dress' cause readers to automatically draw the inference concerning 'clothes.' In contrast Haviland and Clark (1974) found that some inferences took a small but significant amount of time during reading. In their experiment they used pairs of sentences like the following:

(6a) Harry took the picnic things from the trunk.
(6b) The beer was warm.

(7a) Harry took the beer from the trunk.
(7b) The beer was warm.

They found that sentences like (6b) took longer to read than (7b), because the readers had to make the backward inference that the beer was part of the picnic supplies.

Researchers wished to determine the limits on the kinds of inferences which are made immediately and automatically and which are made later. Clearly, there must be limits on the amount of background knowledge readers activate and one of the research goals is to understand what these limits are and the cognitive processes which support this. Models have been developed that attempt to specify the way inferences are made and the way coherent mental representations are derived from texts. (For a fuller account of inferencing *see* **Coherence: Psycholinguistic Approach**).

Although there are significant differences between the models, there are several agreed features of how discourse processing operates in terms of the way readers update their mental representations of a text, the way some information is held in the foreground of processing while others is background, and the way that certain inferences are drawn automatically to maintain a coherent account of the text. For details of the various models of discourse processing that have been proposed (*see* **Discourse Processing**).

Reading as a Developmental and Educational Process

Before an interpretation of a written text can be made, the words on the page have to be read. Although apparently effortless for the skilled adult reader, the processes of identifying letters, recognizing words, and thus distinguishing the meaning conveyed in even a simple sentence, are complex. One key to unraveling how this is accomplished has been to study readers' eye movements. These have been shown to be very systematic and to consist of three main types: short forward movements of around 6–9 letters called saccades, which last on average 20–50 ms; backward movements called regressions; and the pauses or fixations when the readers' eyes rest on a word for around 250 ms.

Rayner and colleagues have shown there are consistent patterns in these movements (see e.g., Rayner, 1998). When reading more difficult texts, fixations grow longer, saccades grow shorter and regressions become more frequent. Even within a single text, readers will spend more time fixating relatively uncommon words compared to familiar ones. Eye movements are designed to keep the middle of our visual field, where our vision is best, aimed at new areas of interest. However, we are able to distinguish quite a lot of information within a single fixation. A skilled adult reader of English will usually be able to identify 15 letters to the right of the fixation point but only three or four letters to the left.

When children begin to learn to read they fixate words for longer than skilled readers. They also have a shorter perceptual span than adults, which means that in a single fixation they are able to identify fewer letters. Gradually these patterns approach those of adults as reading proficiency increases. The typical English reader's asymmetric perceptual span starts to appear in most young readers within a year of starting to learn to read. This seems to reflect the left-to-right nature of reading English, as readers of Hebrew, which is read right to left, show the opposite pattern in their perceptual spans. Like much of psycholinguistics, the study of skilled reading has largely been the study of skilled English reading but more recently interesting studies have been conducted on other languages, notably on nonalphabetic scripts such as Chinese and Japanese. (*see* **Reading Processes in Adults** for more details).

Visual Word Recognition

One aspect of the reading process has received a great deal of research attention in psycholinguistics: the process of visual word recognition. A whole set of phenomena have been identified in the processes of recognizing written words. These include the process of priming. Words are recognized more quickly if they have been read previously. This is known as repetition priming. Words are also recognized more quickly if a word of similar meaning has just been presented, so 'butter' is recognized more quickly if 'bread' has just been read. This is known as semantic priming. Priming can also occur between words which do not seem to have a direct semantic relationship, such as 'music' and 'kidney,' which are linked by an intermediate word 'organ,' which was not presented to the readers.

Other phenomena which researchers have identified in word recognition include the fact that words which are common in the language, such as 'road,' are recognized more quickly than similar words that are less common, such as 'rend.' This word frequency effect is a strong influence on recognition speeds, with even fairly small differences in word frequency influencing reaction times. It is the most robust effect in studies of word recognition, appearing in many different studies using a wide variety of research methods.

These and other phenomena about how words are recognized have been used to develop a variety of models of word recognition. These can be grouped into families of related accounts. Some of the proposed models are direct access models, where perceptual information goes straight to feature counters or units. Other accounts propose that perceptual information is used to trigger a search through the mental lexicon, that is, the stored representation of all the words known to the reader.

One of the best-known serial search models was proposed by Forster (1976). In this account, the perceptual input is used to build a representation of the word to be recognized which is then checked in two stages by comparisons with a series of access files, which are analogous to the cards in a library index system. Once an input string is matched to an access file it is then linked to the master files, analogous to the books on the shelf, which contain the full lexical entries for each word. The files are organized to speed up the process of word recognition, with groups of files being arranged in bins that contain similar words. Within a bin, files are organized by word frequency. The details of the model are described to account for the observed features of the recognition process. For example, there are cross-references between master files that would support semantic priming. Despite these features it is not clear that this kind of serial search model can convincingly account for the speed with which words are recognized. It has also been criticized as being based on a rather dated analogy with the cognitive system as a digital computer rather than as a neural system.

Very recently, however, Murray and Forster (2004) have published an account of a series of word recognition studies that are claimed to show strong support for the serial search model. They claim that the structure within bins, notably the **rank** ordering of words within a bin in terms of frequency, accounts for the pattern of experimental results on word frequency effects more parsimoniously than alternative direct access models.

More popular models of word recognition involve direct access from the sensory information to the lexical units. One of the most influential of the early accounts of this sort was the Logogen Model (Morton, 1969). In this model, perceptual information, either from visual or auditory analysis, feeds directly into the logogen system. This consists of a series of units, logogens, which represent known words. Logogens act as feature counters and when a logogen has accumulated sufficient evidence to reach a threshold it fires. A word then becomes available to the output buffer, is recognized, and can be articulated. The logogens receive input from the cognitive system as well as the sensory input routes and so the resting threshold of the logogen can be varied by, for example, prior experience of a word, or from the sentence or discourse context. Common words with which an individual has had a lot of prior experience will have a lower threshold and will fire with less sensory input and hence be recognized more quickly. Similarly, words that are highly predictable from context will also have their thresholds raised and so will be recognized rapidly.

The Logogen Model has been used as the basis of computational models of word recognition, notably the interaction activation model (IAM) proposed by McLelland and Rumelhart (1981). This was one of the early connectionist or PDP accounts of language processing. Researchers were trying to produce computational models of language processes which were inspired by what was known about the structure and processes of the human brain (*see* **Cognitive Science: Overview**). In word recognition models, this meant that the feature units were simple units with dense series of interconnections, which processed by sending many messages in parallel. The IAM consisted of three layers of units, corresponding at the lowest level to the visual features of letters, then at the next layer to individual letters, then at the top layer to complete words, all with dense arrays of interconnections between them, which like nerve pathways could be excitatory or inhibitory. The model was able to account for a wide range of experimental observations on word recognition. (For more details *see* **Word Recognition, Written**.)

Models of this type have become very widely used in psycholinguistics to explore a wide variety of language processes. This trend towards testing computational models against the existing experimental literature in a way that can link to new insights about the neural processes involved in language and cognitive processes is a popular approach in current psycholinguistics.

Dialogue and Gesture

In parallel to this concern with understanding the cognitive processes and the possible neurological architecture involved in language processing has been a growing interest in language processing in context. This means not only the growth in studies of how extended texts are processed but also a developing field of psycholinguistic studies of interactive language use. These studies have focused on language in its natural setting, interactive dialogue.

Dialogue represents the most ubiquitous form of language use. As young children, we learn to use language through dialogues with our parents and caregivers. Even educated adults spend a great deal of their time in conversations with family, friends, colleagues. Spoken dialogue is still used to obtain many forms of goods and services. In the many nonliterate cultures in the world dialogue is the main or only form of linguistic interaction.

Yet till recently dialogue received rather little research attention in psycholinguistics. One of the challenges for psycholinguists who wished to study dialogue was to derive methods of exploring the phenomenon which would produce testable and generalizable research questions and findings. One experimental method which has been used to allow the study of comparable dialogues from many pairs of speakers is the referential communication paradigm, developed by Krauss and Weinheimer (1964). In this, pairs of speakers are presented with an array of cards depicting abstract shapes. The speakers have to interact to determine which card a speaker is referring to at a given point in the dialogue. These early studies revealed key aspects of dialogue, including the way that over the course of a dialogue, the lengths of descriptions of even complex and abstract stimuli become much shorter and more concise. The interaction between the speakers was found to be crucial in this process. If this interaction was disrupted the speakers were not able to reduce their descriptions to the same extent.

Later studies on referential dialogues were conducted by Clark and colleagues. In one influential paper, Clark and Wilkes-Gibbs (1986) developed a collaborative model of dialogue. From studying many pairs of speakers engaged in dialogues, they highlighted the way speakers work together through their contributions to dialogue to arrive at a shared and mutual way of referring to things they wish to discuss. This means that the speaker and the listener both take responsibility for assuring that what has been said is mutually understood or 'grounded' before the dialogue proceeds.

In this view of dialogue, speakers follow the principle of collaborative effort and try to minimize their overall effort in arriving at an agreed description. This is done iteratively over a number of turns of speaking, often with each speaker contributing part of the description. This process makes use of the opportunities and limitations of spoken dialogue in a way that highlights its differences from written language. In text, the writer can take as much time as is needed to produce a description that she thinks the reader will understand. In dialogue, however, the speaker is under time pressure, as conversation rarely allows long gaps in which to plan an utterance. The speaker therefore has to produce an immediate description and may not be able to retrieve the most appropriate description or judge what way of describing a referent will be interpretable by the listener. The advantage of dialogue, however, is that the speaker and the listener can work together to refine or clarify the speaker's initial description till they both share a common understanding.

The highly collaborative nature of this process is reinforced in two key studies by Clark and colleagues. Clark and Schaefer (1987) showed how contributing to dialogue is characterized by two phases. Each time a speaker wishes to make a contribution, they produce a stretch of speech that is the content of what they wish to contribute. This is called the presentation phase. They then require an acceptance phase, in which their listener gives evidence of understanding the previous contribution. So the dialogue involves two activities: content specification and grounding, that is attempting to ensure that both speakers understand the content sufficiently for their current conversational purposes.

The importance of the ability to actively contribute to the dialogue interaction was demonstrated in a study Schober and Clark (1989). Using the referential paradigm they had pairs of speakers complete the task. An additional participant was included in each interaction who overheard everything that was said but did not take part in the dialogue. This overhearer had a much harder time trying to identify the intended referents of the descriptions than the conversational participant, despite having the same pictures and having heard everything that was said. The suggested explanation is that the descriptions were not grounded for the overhearers. They had no chance to collaborate and ensure that they understood each description as it emerged during the dialogue.

So there is clear evidence that dialogue is an interactive and collaborative process that involves speakers and listeners attempting to cooperate and achieve mutual understanding. The detailed mechanisms that underpin these general processes of adaptation to the interlocutor are now the focus of a good deal of psycholinguistic study. There is some controversy over the extent to which speakers are able to adjust and adapt their output to their listener's needs. In terms of the forms of referring expressions chosen by speakers there is evidence of adjustment to the listener's general level of knowledge of the domain. When it comes to adjusting the intelligibility of their articulation, speakers seem to be largely egocentric. They reduce the clarity of their word production in terms of what is familiar to them as speakers rather than modeling their listeners' needs. The time course of adaptation is also debated, with some studies showing speakers initially produce utterances from their own perspective but later monitor their listeners and adapt. Other studies show listener adaptations from the start of speaking (for more details *see* **Dialogue and Interaction**).

As language processing is studied in more natural contexts of use, it becomes clear that speakers and listeners do not just communicate using the verbal channel. Visual signals from the mouth, face, hands, and eyes are all important features of communication. Researchers have begun to explore the way the visual channel is used by speakers and listeners and the relationship between verbal and visual signals.

More generally, in a dialogue, what we say, how much we say, and even the clarity of the way the words will be spoken have all been shown to change when speakers do or do not have access to visual signals. So speakers who can see one another need to say less to complete a task, use more gestures, can exchange turns of speaking more smoothly, and articulate their words less than when they cannot see one another.

The important role of visual signals in the perception of speech and how these are integrated with acoustic information is a fascinating research area. This was highlighted in a seminal study by McGurk and MacDonald (1976). They demonstrated that when a listener hears a phoneme such as 'ba' while watching a face mouthing 'ga,' the sound which is heard is a fusion 'da.' This is a powerful illusion that occurs even with knowledgeable listeners/viewers and has been demonstrated with young babies (*see* **Audio-visual Speech Processing**).

The role of visual signals in the production and comprehension of more extended stretches of discourse has also been the subject of considerable study. From studies of conversation and storytelling, the important role of gestures and their relationship to the accompanying speech has been established. For some kinds of gestures there is a close temporal relationship with the accompanying speech. Listeners also seem to fuse information presented visually and verbally. If they are told a story by a speaker who uses speech and gesture and are then asked to retell the story later, information originally presented by gesture, such as the speed of an action or the manner of leaving, is often relayed in speech and vice versa (for more information **Gesture and Communication**).

Future Directions in Psycholinguistics

Several trends seem apparent in psycholinguistics. Some of these seem to be the result of improvements and developments in the research methods available to psycholinguistics. One is an increased interest in detailed investigations of language in richer, more naturalistic, contexts. New research techniques such as improved methods of tracking a speaker or listener's eye movements mean, for example, that studies of dialogue, or the relationship of a speaker's production to the surrounding context, can be studied with the precision that used to be only possible in studies of isolated word recognition or sentence processing.

Improvements in the ease and accessibility of various brain-imaging techniques mean that these are being used not only as contribution to our understanding of the neural substrate of different language processes. Techniques such as ERP (event-related brain potentials) can now be used more and more as means to explore the precise time course of language processing. Newer techniques such as MEG (magneto-encephalography) are beginning to offer psycholinguists not just good information about the temporal patterns of language processing but also detailed information about the location of associated brain activity. (For more information *see* **Psycholinguistic Research Methods**.) The growing interest in the neural substrates which support language processing has received a major boost from the development of these new forms of brain imaging.

The interest in how to build neurologically plausible models was of course one of the drivers behind the expansion over the last 20 years in connectionist models of language. These have made a major contribution to our understanding of how a wide variety of language processes, such as spoken or written word recognition, might operate and be learned. The challenge in the future will be to see whether connectionist models can be implemented for more extensive language processing, such as text comprehension or contribution to dialogues. The way such models can or cannot be scaled up to simulate more complex language processing will be one of the key challenges for the next few years.

In the future psycholinguistics will also need to address its undoubted Anglocentric bias. The vast majority of studies of language processing are in fact studies of English language processing. In a number of areas a few studies are emerging which consider other languages but this effort needs to be greatly increased. Over the last 50 years psycholinguistics has expanded dramatically and made considerable progress in understanding a wide variety of language processes. With new research techniques and a more balanced research portfolio in terms of the languages studied and the research efforts applied to production as well as comprehension, spoken as well as written language, future progress seems assured.

See also: Audio-visual Speech Processing; Cognitive Science: Overview; Coherence: Psycholinguistic Approach; Dialogue and Interaction; Discourse Processing; Gesture and Communication; Pauses and Hesitations: Psycholinguistic Approach; Psycholinguistic Research Methods; Psycholinguistics: History; Reading Processes in Adults; Sentence Processing; Speech Errors: Psycholinguistic Approach; Speech Production; Speech Recognition: Psychology Approaches; Word Recognition, Written.

Bibliography

Bailey K & Ferreira F (2003). 'Disfluencies affect the parsing of garden-path sentences.' *Journal of Memory and Language 49,* 183–200.

Bartlett F C (1932). *Remembering: a study in experimental and social psychology.* Cambridge: Cambridge University Press.

Bever T G (1970). 'The cognitive basis for linguistic structure.' In Hayes J R (ed.) *Cognition and the development of language.* New York: John Wiley. 270–362.

Chomsky N (1957). *Syntactic structures.* The Hague: Mouton.

Clark H H & Schaefer E (1989). 'Contributing to discourse.' *Cognitive Science 13,* 259–294.

Clark H H & Wilkes-Gibbs D (1986). 'Referring as a collaborative process.' *Cognition 22,* 1–39.

Clifton C, Traxler M & Mohamed M (2003). 'The use of thematic role information in parsing: syntactic processing autonomy revisited.' *Journal of Memory and Language 49,* 317–334.

Ferreira F & Clifton C (1986). 'The independence of syntactic processing.' *Journal of Memory and Language 25,* 348–368.

Forster K I (1976). 'Accessing the mental lexicon.' In Wales R W E (ed.) *New approaches to language mechanisms.* Amsterdam: North Holland. 257–287.

Frazier L (1979). *On comprehending sentences: syntactic parsing strategies.* West Bend: Indiana Universities Linguistic Club.

Haviland S E & Clark H H (1974). 'What's new? Acquiring new information as a process in comprehension.' *Journal of Verbal Learning and Verbal Behavior 13,* 512–521.

Jusczyk P & Luce P (2002). 'Speech perception and spoken word recognition: past and present.' *Ear and Hearing 23,* 2–40.

Krauss R & Weinheimer S (1964). 'Changes in reference phrases as a function of frequency of use in social interactions: a preliminary study.' *Psychonomic Science 1,* 113–114.

MacDonald M, Pearlmutter N & Seidenberg M (1994). 'Lexical nature of syntactic ambiguity resolution.' *Psychological Review 101,* 676–703.

Marslen-Wilson W (1989). 'Access and integration: projecting sound onto meaning.' In Marslen-Wilson W (ed.) *Lexical access and representation.* Cambridge, MA: Bradford. 3–24.

Marslen-Wilson W & Welch A (1978). 'Processing interactions and lexical access during word recognition in continuous speech.' *Cognitive Psychology 10,* 29–63.

McClelland J & Elman J (1986). 'The TRACE model of speech perception.' *Cognitive Psychology 18,* 1–86.

McClelland J & Rumelhart D (1981). 'An interactive activation model of context effects in letter perception 1: An account of the basic findings.' *Psychological Review 88,* 375–407.

McGurk H & MacDonald J (1976). 'Hearing lips and seeing voices.' *Nature 264,* 746–748.

Miller G & Isard S (1963). 'Some perceptual consequences of linguistic rules.' *Journal of Verbal Learning and Verbal Behavior 2,* 217–228.

Morton J (1969). 'Interaction of information in word recognition.' *Psychological Review 76,* 165–178.

Murray W & Forster K (2004). 'Serial mechanisms in lexical access: the rank hypothesis.' *Psychological Review 111,* 721–756.

Rayner K (1998). 'Eye movements in reading and information processing: 20 years of research.' *Psychological Bulletin 124,* 372–422.

Sanford A J & Garrod S (1981). *Understanding written language*. Chichester, England: John Wiley.
Schober M & Clark H H (1989). 'Understanding by addressees and overhearers.' *Cognitive Psychology 21*, 211–232.
Trueswell J, Tannenhaus M & Garnsey S (1994). 'Semantic influences on parsing: use of thematic role information in syntactic disambiguation.' *Journal of Memory and Language 33*, 285–318.