Psycholinguistic Research Methods

S Garrod, University of Glasgow, Glasgow, UK

Psycholinguistics aims to uncover the mental representations and processes through which people produce and understand language, and it uses a wide range of techniques to do this. The preferred psycholinguistic method is to carry out a controlled experiment. This means that the researcher manipulates an independent linguistic variable to control some aspect of language processing and then measures the effect of the manipulation on a dependent variable of interest. For example, consider an experiment aimed at discovering the influence of the age of acquisition of a word (independent variable) on the time it takes an adult to identify that word (dependent variable). The psycholinguist would construct two lists of words, one containing words learned early in life, the other containing words learned later in life, and then measure the effect of age of acquisition (the independent variable) on the time to identify the word (the dependent variable). To make sure that the independent variable is the real cause of any effect observed in the dependent variable, the experimenter needs to carefully control that variable. For example, in the above experiment it is important to make sure that age of acquisition is not confounded with the length of the word or its citation frequency, because we know that longer words and low frequency words take longer to read.

Once the experiment has been run with a sufficient number of participants and a sufficient range of linguistic materials, the data is analyzed statistically. This usually involves taking the average of the values of the dependent variable (response latencies) for each value of the independent variable (each list of words acquired at different ages) and establishing whether the differences associated with the different values of the independent variable are statistically reliable. It is now standard in psycholinguistics experiments to test that the effects hold true both across the range of participants used in the experiment and across the range of linguistic materials used in the experiment.

Because the psycholinguist is interested in the dynamics of language processing, an important distinction is drawn between **on-line** techniques, which measure variables that tap into language processing as it happens, and **off-line** techniques, which measure variables related to the subsequent outcomes of processing. In practice, on-line and off-line techniques compliment each other, with off-line techniques used to determine the outcome of interpretation and on-line techniques used to determine its time course.

Another major distinction between psycholinguistic techniques relates to the nature of the dependent variables they measure. Some experiments have behavioral dependent variables, such as those associated with a reader's eye movements while reading, others have neurophysiological dependent variables such as those associated with electrical brain activity produced while listening to a sentence. The article first considers behavioral methods and then related neurophysiological methods. It starts with methods for the study of spoken and written language comprehension, and then methods for studying language production and dialogue.

Behavioral Methods

Common Assumptions Underlying Behavioral Methods

Although there is a wide range of behavioral psycholinguistic methods, most depend upon the same basic assumptions. One important assumption concerns how measurements of the time to carry out a task relate to inferences about complexity of processing. Whether the timed response be an eye movement or the time to answer 'yes' or 'no' to a question, it is assumed that the complexity of the mental process is reflected in the response latency. For example, if a reader takes longer to read the fragment "... raced past the barn fell" when it is part of the sentence "the horse raced past the barn fell" than the sentence "the horse that was raced past the barn fell," this is taken to reflect greater complexity in the syntactic analysis of the former than the latter. This assumption is used when interpreting results from most on-line techniques.

A more sophisticated timing methodology investigates the trade-off between speed to respond and accuracy of that response. This is called a **speed accuracy trade-off** (**SAT**) method. With SAT, participants are required to make some response to a linguistic stimulus as soon as they hear a tone, presented at different intervals after the presentation of the stimulus. When the interval is short, participants tend to make many errors, and when sufficiently long, they become completely accurate. So plotting response latency and accuracy across the range of tone intervals gives an unbiased measure of the rate at which the task can be carried out. SAT techniques have been used to assess the rate at which different kinds of lexical, syntactic, and semantic processing occur. Although it is considered a particularly refined technique for establishing processing rates, it has the disadvantage that many thousands of trials have to be run to produce a clear SAT profile, and this will mean that there will have to be many examples of each kind of material used. So there are only a limited number of issues that can easily be investigated with SAT.

Another important assumption underlying many behavioral methods concerns the interpretation of priming effects. **Priming** techniques measure the effect of having previously processed a **prime** item on the subsequent processing of a **target** item. The prime might be a word with a particular form or meaning or it might even be a whole sentence with a particular syntactic structure. The rationale behind priming techniques is that any influence of prime on the subsequent processing of the target must reflect some relationship between the mental representations of prime and target items. Typically, when a prime boosts the interpretation of the target (e.g., by speeding up the recognition of the target item), this is taken as evidence that there is something in common between the representation of target and prime. For example, if a person is quicker to decide that *giraffe* is a word having just read the word *tiger* than having just read the word *timer*, then it is assumed that this is due to the closer semantic relationship between *tiger-giraffe* than between *timer-giraffe*. On the hand, when it hinders the interpretation of the target – this is called **negative priming** as opposed to **positive priming** – this is taken to reflect some conflict between the representations of prime and target. Priming techniques are particularly useful for establishing the relationship between different linguistic representations used during processing and are widely used in conjunction with a number of behavioral or neurophysiological measures.

It is beyond the scope of this article to describe all the behavioral techniques that have been used in psycholinguistics. Instead, the article concentrates on techniques that have had a major impact on the field.

Behavioral Methods for Spoken Language Comprehension

Cross-modal Priming A common priming technique used to tap into spoken language comprehension is what is called **cross-modal priming**. In cross-modal priming the prime item is usually a word embedded in a spoken sentence or text used to prime a target item presented in written form, which is why it is called cross-modal. Consider the problem of working out how listeners resolve ambiguous words. Cross-modal priming can indicate the immediate interpretation of an ambiguous word, such as *bug*, in contexts that promote either one or other meaning of the word (e.g., 'insect' or 'listening device'). As participants listen to *bug* in the different contexts, they are presented with a written word (ANT or SPY) or a nonword (AST) and have to decide as quickly as possible whether the target is a word or not (this is called **lexical decision**). The question of interest is whether lexical decision of the targets ANT and SPY is boosted by hearing the prime *bug* in the different contexts. It turns out that when the visual target is presented immediately after hearing the prime in either context it promotes lexical decision for both targets. So this indicates that both interpretations ('insect' and 'listening device') are immediately activated irrespective of the disambiguating context. However, if the visual target is presented at a slightly later point (about 200 ms after the prime), only the target related to the contextually appropriate interpretation (i.e., ANT for 'insect' or SPY for 'listening device') is

primed. Cross-modal priming therefore indicates that all meanings of ambiguous words are accessed immediately on encounter, but then only the contextually appropriate meanings are retained as comprehension proceeds.

Cross-modal priming has been used to investigate many aspects of spoken language comprehension. These include influences of prior discourse context on lexical processing, resolution of anaphoric references (e.g., pronouns), and morphological analysis. The technique has the advantage that it can tap into spoken language comprehension as it occurs and often uncovers aspects of processing of which the subject is completely unaware, such as in the ambiguity experiment described above.

The Gating Technique Another technique for establishing when listeners interpret words in relation to information in the speech stream is **gating**. The gating procedure involves presenting increasingly long fragments of speech and measuring when listeners can interpret the speech appropriately. For example, if you want to know at what point a listener can accurately recognize the word *cathedral*, you can present them with the spoken fragments /ca/, /cath/, /cathe/ , /cathed/ ... and record the listener's judgements for each fragment. The shortest fragment that can be correctly identified defines the point at which there is sufficient information in the speech for identification. This technique can be used to estimate the earliest point at which the listener could identify a word and so make it possible to test whether this predicts the recognition time for that spoken word.

The Visual World Paradigm One of the most effective on-line measures is **eyetracking** (recording the precise pattern of eye fixations during comprehension). Because visual attention is strongly controlled by where a person is currently looking, eyetracking can be used to indicate what a person is attending to at any point during comprehension and for how long they attend to it. The technique can either be used for reading research (see below), in which case the focus of attention corresponds to the words being looked at any time or it can be used to measure which part of a scene a participant attends to as they interpret spoken utterances about that scene. The latter technique is usually referred to as the **visual world paradigm**. A classic example of the use of this technique is in determining how listeners deal with syntactic ambiguities that arise during comprehension. For example, participants listen to instructions that are initially consistent with two syntactic analyses while they view a scene containing a small number of objects. They might be asked to *Put the frog on the napkin in the box* when there is either one or two frogs in the scene. Analyses of eye movements demonstrate that viewers look at the frog which is on a napkin more if there is also a frog that is not on a napkin than if there is no other frog present. From this it can be inferred that the visual context drives syntactic disambiguation (i.e., it supports the reading in which *on the napkin* syntactically modifies *frog* rather than being a syntactic argument of *put*) at an early stage in processing.

The visual world paradigm has been used to investigate a wide range of issues in spoken language comprehension. These include lexical access, resolution of anaphoric pronouns, development of strategies for semantic and syntactic processing, language development, and even language processing in dialogue. It offers a precise indication of when listeners integrate information from a linguistic utterance with that in the visual world and tends to show that comprehension is both incremental and immediate in relation to most levels of linguistic analysis.

Behavioral Methods for Written Language Comprehension

There are a variety of techniques that tap into the time course of written language comprehension.

Self-paced Reading One class of techniques is **self-paced reading**. The reader determines the rate at which written material is presented and the experimenter records the rate of presentation. A reader might be required to pace himself or herself sentence by sentence, phrase by phrase, or word by word. For example, in the word-by-word procedure, a word is presented, and as soon as the reader has understood it, he or she presses a key to trigger presentation of the next word. The sequence is then repeated until all the text has been read and the time to read each word is recorded. This kind of technique has been used to study syntactic analysis, discourse comprehension processes and in particular resolution of anaphors. It gives a good indication of when a reader encounters difficulty in comprehension, but is limited according to the size of linguistic unit being presented. Whereas larger units such as whole sentences can be read at a normal rate during self-paced reading, smaller units like words tend to be read much more slowly in self-paced reading tasks. Hence, when the technique has high on-line resolution (e.g., when word by word), it also interferes most with the normal reading process. This is not a problem with the eyetracking technique described below.

Rapid Serial Visual Presentation (RSVP) A slightly different technique for presenting written language

uses what is called **rapid serial visual presentation** (**RSVP**). With RSVP, readers see sequences of words in the center of a computer screen presented at a fixed fast rate. The experimenter then has the reader carry out an additional task, such as identifying a word in the sequence or trying to recall the sequence of words, which indicates how comprehension is limited by the rate of presentation. This technique has been used to investigate lexical, semantic, and syntactic processing. Like word-by-word self-paced reading, RSVP may well interfere with normal language processing.

Eyetracking During Reading The least interfering on-line behavioral technique for written language comprehension is eyetracking. During reading, the eye moves in a systematic way. There are brief **fixations** in which gaze stays on the same letter interspersed with fast movements called **saccades** during which the gaze moves to another letter or word of the text. For a skilled reader, 9 out of 10 saccades move the gaze from left to right to sample new material from the text, whereas 1 out of 10 saccades return the point of gaze to previously read material (these are called **regressions**). The duration of fixations and the length and direction of saccades (i.e., forward or backward movement of the gaze) directly reflect the ease or difficulty of the reading process. Furthermore, they indicate the precise word in the text that is causing reading difficulty because attention is only given to the word currently fixated.

The limited span of attention during reading can be demonstrated using the **moving window** technique in which a computer program controls dynamically the window of text presented to the reader as a function of where they are fixating. For example, with an asymmetric 12-letter window, the 4 letters to left and the 8 letters to the right of where the reader is fixating will be displayed as normal, whereas all the remaining text will be converted into random letters. The window of text together with its surround of random letters then changes as the point of fixation changes. One can reduce the size and form of the text window and measure when it begins to affect reading rate. It turns out that normal reading is quite possible when the window only contains the word currently fixated plus the first three letters of the next word on the line. However, there is a proviso that the material around the window must retain the spaces between the words in the original text. When the window arrangement is reversed so that the window contains random letters and the surround contains the normal text, readers encounter difficulty. With a reverse window of only 11 letters in width, reading becomes almost impossible.

Moving window studies indicate that readers only take in information from a limited region of text at any time during reading. This means that any extra time spent fixating the region must reflect processing difficulty associated with that region of text or previously fixated regions of text not completely processed but still held in memory.

Eyetracking has been used to study a wide range of linguistic processes, including lexical access, resolving lexical ambiguities, syntactic analysis, and various discourse processing phenomena, such as anaphora resolution. It is particularly effective in determining precisely when the reader makes a decision about some aspect of the linguistic input during sentence or discourse processing. For example, when presented with the sentence *We like the city that the author wrote unceasingly and with great dedication about* readers spend longer fixating the verb *wrote* as compared with the same verb in the fragment *We like the book that the author wrote unceasingly and with great dedication about*. This shows that readers immediately attempt to integrate each word of the sentence with the prior discourse and hence they detect the temporary anomaly produced by *wrote* in the first but not in the second sentence.

Eyetracking is a particularly effective technique because it does not interfere with the normal process of reading. A similar claim is made for some neurophysiological techniques such as ERP, described later.

Behavioral Methods for Spoken Language Production

Until the last decade, language production was not a central topic in psycholinguistics. This was partly because researchers did not have on-line methods for studying production in properly controlled experiments. This meant that the pioneering research on language production depended on off-line techniques, such as the study of speech errors. However, more recent work in language production has been influenced by on-line techniques. We consider both kinds of technique below.

Analysis of Speech Errors Speech error data has been used to draw many interesting conclusions about the nature of language production. For example, it was observed that substitution errors (e.g., saying *if **I** was done to that* rather than *If that was done to **me***) tend to always involve the same syntactic classes (e.g, the pronouns *me* and *that*). Also, as the example shows, the lexical substitution does not always involve a syntactic substitution. Otherwise, the error would have produced *If **me** was done to that*.

Such findings provide evidence that in production, words are chosen before they are strung together to make up an utterance and that this occurs before the words are marked for syntactic case.

Speech errors have been used to argue for an overall organization of speech production into separate message planning, lexical and grammatical assembly, and phonological processing components. There have also been some attempts to develop techniques to elicit speech errors experimentally by priming the errors in a similar fashion to that used in tongue twisters. Many of the conclusions drawn from the analysis of speech errors have been supported by results from more recent on-line techniques such as picture naming.

Picture Naming A particularly influential on-line technique for studying language production is to measure the time it takes participants to name pictures of objects or events. **Picture naming** is commonly combined with some form of priming task (see next section). It can also be combined with other techniques such as eyetracking to give a more precise indication of how words are accessed during language production. For example, a researcher might want to know whether speakers wait until they have all the words available before they start to articulate the first word in an utterance. To find out, the researcher could have participants name two pictured objects *A* and *B* in a phrase of the form *A and B*. If the time to start articulating the name of *A* is unaffected by the time to access the name for *B* as indicated by the time to articulate *B* in isolation, then the researcher can conclude that the speech articulation process begins before both word forms have been accessed. In fact, the evidence supports this conclusion.

Priming Techniques in Language Production An important issue in language production concerns the extent to which utterances are formulated incrementally one unit at a time according to different levels of representation (semantic, syntactic, and phonological). Priming techniques have been used to address this and other related questions. For example, at the phonological level words could be assembled for articulation either as complete packages or incrementally as sequences of distinct phonological units. A so-called **implicit priming** technique has been developed to test this. The participant has to learn sets of pairs of words, such that when given the first word in the pair he or she names the second word as quickly as possible. Crucially, in one condition the second words always share the same first syllable (e.g., single-loner, place-local, fruit-lotus), whereas in the other condition they do not (e.g., single-loner, signal-beacon, captain-major). The question is whether participants can use the implicit prime of the shared first syllable to speed up articulation. It turns out that they can, and more interestingly, the implicit priming only works for the first syllable in the word. When given a comparable list in which the second syllable of the second word is shared (e.g., single-murder, place-ponder, fruit-boulder), there is no articulatory benefit.

There are also techniques for studying priming at the syntactic level in production which use a variant of the picture naming procedure. A typical study might involve participants describing a sequence of pictured events using a verb indicated at the bottom of each picture. Interleaved between these descriptions, the participant checks descriptions they hear against another series of pictured events. By using ditransitive verbs such as *give*, which take either prepositional objects (*give the picture to Mary)* or double objects (*give Mary the picture*), it is possible to study syntactic priming independent of semantic priming. The question is whether having just heard *The sailor gave the banana to the nun* participants are more likely to describe their next picture as *The clown handed the book to the pirate* than *The clown handed the pirate the book*? It turns out that they are.

Behavioral Methods for the Study of Dialogue

As with language production, it is only quite recently that psycholinguists have begun the experimental investigation of language processing during dialogue. Again there is what has sometimes been called the problem of exuberant responses. Because dialogue is inherently spontaneous, how can the experimenter exert the control required for a sound experiment?

One way around this problem is to set up a task that controls what interlocutors can talk about. One such task is the **referential communication task**. One participant is required to describe a series of arbitrary visual patterns such that their partner who is able to reply can identify each pattern from the set that they have been given. This technique has been used to investigate how feedback from the listener affects the nature of subsequent references made by an interlocutor. More complex dialogue tasks have had conversational partners describe routes on a map or positions in a maze as part of some other cooperative activity. More recently, researchers have started to record interlocutors patterns of eye movements while carrying out some version of the referential communication task. In this way, it is possible to combine aspects of the visual world paradigm with those of the interactive referential communication task.

Neurophysiological Methods – ERP, fMRI, and MEG

Techniques for measuring the neurophysiological correlates of language processing have recently increased the psycholinguist's methodological armory. Three particular measurement techniques have been used. The first measures electrical activity at the scalp using **electro-encephalography** (**EEG**) to produce what are called **event-related brain potentials** (**ERPs**). The second measures changes in brain blood flow associated with neural activity using **functional magnetic resonance imaging** techniques (**fMRI**), and the third measures changes in magnetic fields associated with the electrical activity in the brain using **magneto-encephalography** (**MEG**). Each technique has its own advantages and disadvantages as a psycholinguistic tool.

fMRI signals give precise information about the area in the brain associated with the particular activity, so fMRI has proved useful for neurolinguistic investigation. The disadvantage with the technique for psycholinguistics is that it is not good for establishing the time course of the neural activity. This is because it takes time for the neural activity to produce changes in the blood flow. By contrast, ERPs provide precise information about the time course of the neural activity, but it is difficult to establish the source of the activity. This is because the ERP signal is affected by all sorts of irrelevant factors such as the thickness of the skull or interactions between signals from different areas of the brain. Finally, the most recently developed technique, MEG, offers good localization with similar temporal resolution to ERP. Nevertheless, like ERP signals, MEG signals are only sensitive to activity in neural structures with particular orientations, and neither technique is easy to apply when there is contemporaneous motor activity, such as eye movements or articulation during speech.

The most influential psycholinguistic research to date has tended to use ERP because it is both relatively cheap to run ERP experiments and the technique offers similar temporal resolution to that of behavioral measures such as eye tracking. This article concentrates mainly on the use of ERPs, but it will also say something about the use of MEG.

Event-related Brain Potentials (ERPs)

ERPs are derived from measurements of small changes in voltage at different points across the scalp. They are called **event-related potentials** because they are analyzed in relation to the onset of a triggering event. For instance, to derive an ERP that reflects word identification processes, the experimenter presents a word on a computer screen and measures the changes in scalp voltage from the point at which the word was presented. This process is then repeated over a number of trials. Because the data from each trial will contain irrelevant electrical activity, the experimenter takes the average potential across the set of related trials. In this way, the irrelevant 'noise' information can be filtered out. What remains is an ERP waveform with identifiable peaks and troughs of voltage. These peaks and troughs are taken to reflect the activity of bundles of nerve fibers in particular parts of the brain.

In practice, ERP researchers try to identify the characteristic peaks and troughs and establish how they might relate to concurrent processing. One well-established peak is called the **N400**, which corresponds to a negative component of the wave occurring approximately 400 ms. after presentation of the word that triggers it (it is a peak because ERP researchers conventionally plot negative values upward and positive values downward). The N400 has been associated with processes of conceptual or semantic integration of words into their sentential contexts. For example, when you measure ERPs elicited by the words *eat, drink,* or *cry* in the context "The pizza was too hot to *eat/drink/cry,*" then the N400 is larger for *cry* than *drink* and larger for *drink* than the contextually appropriate *eat*. This pattern is consistent with the idea that the N400 reflects conceptual integration (with *eat* being integrated into the sentence better than *drink*) as well as semantic integration (*drink* is semantically related to *eat* whereas *cry* is not).

Another characteristic trough is called the **P600** (a positive change occurring about 600 ms after the triggering word). Unlike the N400, the P600 has been associated with syntactic integration processes. For example, the word *was* when presented in the ungrammatical sentence *The doctor forced the patient was lying* produces a much larger P600 than *was* in the sentence *The doctor thought the patient was lying*. Because these two wave forms, N400 and P600, seem to reflect different kinds of processing, ERP can be used to establish the precise time course of these different processes. So ERP complements other on-line measures such as eyetracking, which do not differentiate between different kinds of psycholinguistic processes in this way.

ERPs have been used to address a wide range of questions both about early stages of lexical processing and more general syntactic and semantic processes in language comprehension. It is best suited for the study of processes that immediately follow presentation of the triggering stimulus because the ERP signal becomes increasingly noisy over time. This means that it is not such a good technique for studying such things as syntactic re-analysis or integration of a sentence into the discourse context.

Magneto-encephalography (MEG)

MEG has only recently begun to be used in psycholinguistics. MEG has some of the advantages of both fMRI techniques and EEG techniques because it enables precise source localization like fMRI and has fine temporal resolution like ERP. It also complements ERP. Whereas ERP electrical signals can only be picked up from nerve bundles that are in particular orientations with respect to the surface of the brain, MEG magnetic field signals can be picked up from nerve bundles that are orthogonal to those giving ERP signals. For these reasons, many researchers are particularly optimistic about MEG as a psycholinguistic method used together with ERP.

One particularly interesting application has been using MEG to establish the relationship between neural representation and linguistic form. The technique depends upon what has been called **mismatch negativity**. It was observed that as the same items are repeatedly presented to subjects, so the MEG signature associated with their processing is automatically reduced (probably because of neuronal habituation). However, when a new item is presented, the signal returns to normal. This happens irrespective of any behavior on the part of the subject. Mismatch negativity can therefore be used to establish the degree to which different items are processed in the same way. The greater the resumption of the activity (i.e., mismatch negativity), the more different the neurological processing of the new item. In this way, mismatch negativity can be used in a similar fashion to priming techniques to explore the neurological representation of different aspects of a linguistic stimulus.

Summary and Conclusion

Psycholinguistic techniques differ according to the kind of variables measured and the extent to which they tap into language processing as it happens. Behavioral measures, such as eyetracking, and neurophysiological measures, such as ERP, are particularly effective for measuring the time course of language comprehension. For language production studies, picture naming and priming techniques have been especially effective.

See also: Dialogue and Interaction; Evoked Potentials; fMRI Studies of Language; Magnetoencephalography; Psycholinguistics: Overview; Reading Processes in Adults; Speech Errors: Psycholinguistic Approach; Spoken Language Production: Psycholinguistic Approach.

Bibliography

Carreiras M & Clifton C E (eds.) (2004). *The on-line study of sentence comprehension: Eyetracking, ERP and beyond*. Hove: Psychology Press.

Garrod S (1999). 'The challenge of dialogue for theories of language processing.' In Garrod S & Pickering M J (eds.) *Language processing*. Hove: Psychology Press. 389–415.

Levelt W J M (1989). *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.

Osterhout L & Holcomb P J (1995). 'Event-related potentials and language comprehension.' In Rugg M D & Coles M G H (eds.) *Electrophysiology of mind: event-related brain potentials and cognition*. Oxford: Oxford University Press. 171–216.

Pulvermüller F (2001). 'Brain reflections of words and their meanings.' *Trends in Cognitive Sciences 5,* 517–524.

Rayner K & Sereno S C (1994). 'Eye movements in reading: Psycholinguistic studies.' In Gernsbacher M A (ed.) *Handbook of Psycholinguistics*. San Diego: Academic Press. 57–82.

Trueswell J C & Tanenhaus (eds.) (2004). *Approaches to studying world-situated language use: bridging the language-as-product and language-as-action traditions*. Cambridge, MA: MIT Press.