18
Psycholinguistics

Language and cognition

Matthew J. Traxler

18.1 Introduction: Why is psycholinguistics significant?

We are what we do. What psycholinguists do is apply research methods derived mainly from the discipline of psychology to understand human behavior, thought, and neural processes as they relate to language function. Mainly because psycholinguists are a diverse bunch, they bring a diverse set of methods to the task of understanding how humans produce and understand language. Researchers in the field of psycholinguistics include psychologists and linguists (of course, so long as said linguists are interested in mind, brain and behavior), but also computer scientists, anthropologists, sociologists, philosophers, statisticians, historians, literary scholars and a variety of other disciplines. Psycholinguistics, therefore, provides a prototypical example of a multidisciplinary approach to science.

Psycholinguistics is a science, because psycholinguists apply the scientific method (observe, hypothesize, test, repeat), and it is as data-driven as any other scientific enterprise. The study of language in psychology at its inception was characterized by armchair philosophy, but those days are long gone. Beginning students of psycholinguistics are often surprised that it is a science, and that they therefore need to learn how to apply the scientific method to language. This involves understanding how experiments are designed, how data are collected and analyzed, and how the human nervous system works.

Psycholinguists have made numerous contributions to our theoretical understanding of how people represent knowledge about language, how people access and use this knowledge in real time to produce and understand language, how the human species acquired its current range of language abilities, and how neural systems support language function. These contributions are important not only because understanding how language works is a necessary part of understanding how human beings work, but also because there is a close relationship between language disorder and other kinds of mental disorder. This close connection between language function and overall mental function most likely reflects bi-directional causal relationships from language to non-linguistic cognition, and vice versa. When non-linguistic aspects of cognition (such as attention, executive control or working memory) go wrong, aspects of language function can go wrong. When aspects of language function go wrong (as in Specific Language Disorder, dyslexia and aphasia), this has
consequences for the social and emotional functioning of the individual, as well as effects on educational and economic outcomes. Here is an example of each type of causal relation:

Language function in schizophrenia may represent a case where general cognitive deficits lead to problems with the production and understanding of language. Schizophrenia is a mental disorder characterized by problems in language, hallucinations and delusions, disordered social and emotional function, and disordered motivation. In addition, patients with schizophrenia experience physical changes to the structure of the brain that appear to contribute to general problems in cognition. Patients with schizophrenia experience deficits in cognitive control. Cognitive control, roughly speaking, reflects the ability to regulate attention and to select task-appropriate responses when more than one response is available. Deficits in cognitive control may be caused by mis-regulation of neural systems in the frontal lobes as well as disordered functional connectivity between frontal lobe networks and other, more distant parts of the brain. Patients with schizophrenia experience deficits in attention and memory (both working and long-term memory processes) and they also often have dramatic language-related symptoms (which were commented on during the very earliest era of research into schizophrenia by scientists such as Eugen Bleuler). Patients with schizophrenia have difficulty maintaining a coherent topic during conversation. Their responses to conversational prompts often include references to concepts that are only very loosely related to the prompt. They do not typically have problems with more basic aspects of language production, such as articulation (i.e. their speech, unlike the speech of patients with aphasia, is smooth and essentially normal with respect to grammar).

Psycholinguists have developed processing accounts, derived from work on normally functioning individuals, that help us understand how changes in cognitive control lead to some aspects of disordered language production and comprehension in patients with schizophrenia. In particular, these accounts help us understand why patients with schizophrenia have trouble maintaining a topic and why their production is characterized by loose associations. These accounts appeal to the notion of a semantic network, which is a hypothetical mental mechanism that represents your knowledge of what words mean, and how different word meanings relate to one another. In semantic network theory, concepts that are more similar in meaning are represented ‘closer’ together in the network. As difference in meaning increases, ‘distance’ (virtual distance) in the network increases. These representational assumptions help explain why pairs of words that are closely related (e.g. cat–dog, mother–father) are processed more readily than pairs of words that are more distantly related (e.g. lion–stripes; stripes is related to lion via the intermediary tiger; lion is related to tiger, tiger is related to stripes). According to Condray et al. (2003), activation spreads uncontrollably through the semantic network in patients with schizophrenia because of the lack of ‘top-down’ control that is normally exerted by executive control systems. This uncontrolled spread of activation means that more distantly related concepts (e.g. stripes when lion is mentioned) become activated during production and comprehension, even though those concepts are not relevant to the topic. Uncontrolled spread of activation leads to mental representations that are cluttered with irrelevant and unhelpful information. This greatly complicates the process of understanding and producing language. Hypothetically, if executive control were restored in these patients, this would improve their language function, because better executive control processes could regulate the activation of irrelevant concepts (i.e. executive control mechanisms could reduce the activation of irrelevant concepts; clearing the way for relevant information to come to the forefront).

Specific Language Impairment (SLI) provides an example of language dysfunction leading to disabilities in other aspects of life. SLI is defined somewhat differently by
different researchers, but roughly speaking, SLI is diagnosed in individuals whose language function is substantially lower than one would expect based on overall cognitive skills. In SLI, something seems to be wrong with language that is not explainable as being a consequence of some other thought-related problem. This definition excludes individuals with Down syndrome, for example, because language problems in Down syndrome are likely caused by lower overall intellectual functioning. To qualify for a diagnosis of SLI, people must be free of ‘frank neurological defects’ (this excludes people with neurodevelopmental disorders and people with acquired brain damage, which is the most common cause of the aphasic language disorders). There is some evidence that SLI in some individuals results from a genetic mutation on the FOXP2 gene.

People with SLI can have a number of language-related problems at different levels. These may include reduced ability to perceive and differentiate speech sounds (deficits in phonological function). They may also include reduced ability to comprehend and produce tense, aspect and agreement markers on verbs (e.g. the correct first-person, singular form of the verb to walk is walk; the correct third-person, singular form is walks; as in I walk to the store; He walks to the store).

Language problems in SLI have consequences for social, educational and economic outcomes. Children with SLI make fewer attempts to initiate conversational exchanges with other people. Lack of practice with verbal and social exchanges means that children with SLI miss out on opportunities to sharpen their language and social skills. Research on educational settings shows that children with SLI are viewed differently than children with normal language function. Their peers are less likely to identify them as preferred playmates. Mutual preference (i.e. two children both saying that they prefer each other as a playmate) fosters the development of interpersonal relationships in school-aged children. Unfortunately, a child with SLI is less likely to have his or her preferences reciprocated by another child. Hence, children with SLI are likely to feel more socially isolated than other children. Development of reading skill in children with SLI places limits on their economic success later in life. Hypothetically, if one could intervene to improve language function in individuals with SLI, this would lead to improvements in other areas of function.

Understanding language disorder (its causes and consequences) is an important step in designing treatment methods that alleviate suffering. Thus, more and more, psycholinguists are getting interested in translational research. Translational research does not mean studying how you translate from one language to another. Instead, translational research happens when you use findings from basic science to understand some process (such as language comprehension or production) and then use that understanding to get better control over some condition. As implied by the two examples above, there is tremendous need for translational research in the area of language function. This is especially true in the area of language acquisition.

Research indicates that developmental language disorders (language disorders that occur in children that prevent them from achieving the full range of language ability that normally functioning children have) occur for several different reasons. Sometimes causation runs from abnormal environment to abnormal cognitive development to abnormal language function. Children who are abused or neglected tend to experience greater delays in hitting major language milestones (speaking the first word; using words in combination; acquiring grammatical principles or rules). Sometimes causation starts before birth. Children who are born before full term (approximately thirty-seven weeks of gestation) can develop problems with language, presumably because neural and general cognitive development lags behind.
full-term infants. The best estimates suggest that between 6 and 8 percent of children will experience some kind of serious language delay or disorder.

Language disorders produce some clear disadvantages. For instance, children who experience difficulty learning how to produce or understand language do worse in school than children who readily learn to speak and understand. Still worse, children who experience serious delays hitting major language milestones experience major psychological problems at a higher rate than their normally developing peers. These psychological problems include anxiety, depression, deficits in social interaction skills, major behavior problems and juvenile delinquency. These problems make these children more prone to being expelled from school and reduce their opportunities for employment. More than half of children with significant language delays experience serious deficits in behavior and are at great risk for being diagnosed with a psychiatric disorder. For comparison, the rate of similar problems (lifetime prevalence) in the general population is about 10 to 15 percent. On the flip-side, some studies indicate that up to 80 percent of children who receive a psychiatric diagnosis involving behavior control have a previously undiagnosed language deficit.

Even though social, emotional and behavioral problems correlate very highly with language dysfunction, only a small proportion of children receive evaluation and treatment specifically for language problems until something else goes wrong. In the absence of rigorous evaluation, many teachers, parents and other care-givers misunderstand the nature of the child’s difficulty. Children with undiagnosed language problems are oftentimes viewed as merely ‘delinquent,’ ‘depressed’ or ‘aggressive.’

The precise nature of the relationship between language problems and other psychological problems has not been completely worked out. One plausible hypothesis views language as the cause, and behavioral and psychiatric problems as the consequence. Under this hypothesis, children who experience difficulty producing or understanding language may be less able to communicate their thoughts, feelings and desires. This inability to communicate effectively in turn leads to maladaptive patterns of thought and behavior. This hypothesis is supported by the finding that language function in young children predicts the likelihood of behavioral problems during the elementary school years. There may be no simple link between language function and behavioral outcomes, however. Other factors may mediate this relationship. For example, problems producing and understanding language contribute to social isolation, because children with language problems experience difficulty initiating and maintaining social relationships with peers and care-givers. Children who have language problems are rated as less desirable playmates by their peers, and this pattern starts to emerge in preschool. When children are socially isolated, they lose opportunities to learn all of the subtle cues that children pick up during social interactions. Lesser language skills, possibly in combination with poor social interaction skills, can promote maladaptive behavior. When a person has difficulty using language to communicate, they may adopt other strategies, such as engaging in aggressive or risky behavior, in an attempt to get their desires met. Alternatively, there may be a reciprocal, interactive relationship, where language difficulties and behavioral problems reinforce one another. Children who have behavior problems miss opportunities to strengthen their communication skills; and weak communication skills make it more likely that a child will have behavior problems.5

Clearly, what we need are better theories of language function, and better theories of how to intervene to improve language function. This is the goal that many psycholinguists pursue in their research careers. The following sections provide a brief review of major trends in psycholinguistic theory and research over the past half century.
18.2 A short history of psycholinguistic contributions to advances in the study of language and cognition

18.2.1 The first wave: the cognitive revolution

The study of language in psychology extends back to the very beginnings of the discipline. Wilhelm Wundt, who founded the scientific study of psychology, is credited with inventing the tree diagram, a way of representing the syntax of a sentence. The early study of cognition in psychology was based largely on introspection (reflecting on the content of one’s own thoughts). This approach to understanding psychology was supplanted in the first half of the twentieth century by behaviorism. Theorists such as Edward Thorndyke, J.B. Watson and B.F. Skinner rejected subjective impression and introspection as a way of investigating psychological phenomena. They insisted that theories be based on directly observable events, specifically the behaviors that people and animals produced. Appeals to hidden mental events were outlawed.

Behaviorism focused on the way animals learn associations between stimuli (as in classical conditioning) and the way animals learn associations between situations, their own behaviors and the consequences of those behaviors (reward and punishment). Behaviorists also developed descriptions of language that were firmly grounded in the relationship between behavior and reward. According to these accounts, thought was merely subvocal speech; i.e. a very subtle form of implicit language behavior. J.B. Watson viewed ‘thought’ as occurring because the muscles that were involved in articulating speech were moving (very, very subtly), even though no sound waves resulted from those movements. If thought is really just covert behavior, the reasoning went, then a satisfactory theory of behavior could be straightforwardly extended to account for language.

B.F. Skinner developed a theory of language acquisition based on behaviorist principles according to which children learned to speak because their speech behaviors were rewarded by their care-givers. Skinner also had a theory of speech production and comprehension, based on associations. According to this account, languages can be thought of as patterns of associations between adjacent words. Why do we produce some sentences and not others? According to Skinner, it is because the associations between some words are stronger than the associations between other words. We learn these associations because we are exposed to those patterns when we listen to other people speak. These patterns of associations can be captured in formal arithmetic models (called ‘Markov chains’ or ‘transition networks’). According to this kind of account, you are more likely to produce a noun after an adjective than a verb (green leaves is more likely than *green leave), because there is a much stronger association between adjective and noun than there is between adjective and verb (in another version of the account, associations between individual words, rather than classes of words, drive learning). A Markov chain with the right map of associations is capable of generating a very large, if not infinite, number of sentences, all of which would be compatible with the kinds of sentences that real people produce in real exchanges.

Noam Chomsky (along with scientists like George Miller, Steve Kosslyn and Donald Broadbent) is credited with initiating the cognitive revolution, starting in the 1950s. During this period, cognitive theories overturned and supplanted the behaviorist understanding of language acquisition, production and comprehension. While certain aspects of behaviorism remain intact (e.g. the idea that consequences govern behavior), the behaviorist description of language function has been largely if not completely abandoned.
Cognitive scientists working on a broad spectrum of phenomena, including memory, attention, perception and language, appeal to hypothetical mental processes to explain various aspects of human thought and behavior. These appeals to ‘invisible’ mental processes were strictly taboo to followers of radical behaviorism, the dominant paradigm in psychology from the early 1900s to the late 1950s. To the behaviorists, no theory that appealed to unobservable mental processes was acceptable. This meant that concepts like traces of events in long-term memory, internal manipulation of mental representations, and internal control processes, were not legitimate objects of study. Because these mental representations and processes could not be directly observed, they were viewed by the behaviorists as being illegitimate, or at least not worthy of investigation.

Cognitive scientists therefore faced two hurdles, one representing a negative agenda and the other positive. The negative agenda involved finding ways to discredit behaviorist approaches to various psychological phenomena. The positive agenda involved constructing well-supported, or at least plausible, alternative theories of the processes in question. Chomsky provided convincing arguments for the negative agenda – that behaviorism provided incomplete and inaccurate descriptions of language phenomena. Chomsky and the psychologists who took up his cause pursued the positive agenda as well.

Chomsky’s critique pointed out several observable language-related phenomena for which the behaviorist account had no ready explanation. These critiques were grounded in observable aspects of children’s language acquisition and aspects of the behavior of mature speakers. In that way, Chomsky’s theorizing was not vulnerable to the criticism often leveled by behaviorists, that appeals to ‘invisible’ mental processes were equivalent to appeals to ‘spirits’ or ‘invisible forces.’ Two language-related phenomena help illustrate the general point. According to behaviorists, children learn language the same way they learn other behaviors, by trial and error and by the reinforcement of some behaviors but not others. The behaviorist theory of language therefore predicts two things: First, children’s speech will include many errors that do not occur in the adult language. Second, children will more readily learn aspects of language that they are rewarded for producing; they will be less successful learning aspects of language for which they do not receive feedback (in the form of correction or reward). Neither of these predictions is borne out by data from children acquiring their first language.

In terms of errors, children definitely do make errors. They produce utterances that they never hear in ambient language. However, some kinds of errors that are likely on the behaviorist view are never produced by actual children learning their first language. Behaviorist theory does not have a means to explain why some errors occur and others do not, but a cognitive account of language that includes ‘invisible’ mental operations (mental processes that you cannot see directly) can explain those patterns. One classic example involves the way questions are formed. According to cognitive approaches to language, questions are formed by the modification of simple (canonical) underlying mental representations. You can create a question from the underlying form *The little girl is leaving* by moving the verb *is* to the front, as in *Is the little girl leaving?* This mental formalism explains how questions are generated and why there is a clear relationship in meaning between the statement and the question (i.e. they share an underlying representation).

An associationist/behaviorist view of this process might say that you learn to make questions because you notice that the word *is* appears at the beginning of some expressions. That will work only some of the time, however. You cannot make a question out of this statement *The little girl that is talking to Susan is leaving* by moving the first *is* to the front (*Is the little girl that talking to Susan is leaving?*), but you can make a question by moving
the second *Is the little girl that is talking to Susan leaving?*). Therefore, not everything that has *is* at the beginning is OK. However, the behaviorist view has no explanation for why one form of question is OK, while the other is not. If *is* can precede the *little girl* in one question (presumably because *is* has been associated with the *little girl* somehow), then the pattern of association should allow *is* to precede the *little girl* in any other question. If the association is somehow blocked in one question, it should be blocked in all of them.

The cognitive view of language, on the other hand, proposes that the hypothetical mental element main clause and main verb have a different status than the hypothetical mental element relative clause and relative clause verb. The cognitive view proposes that speakers can mentally rearrange the elements in an expression by moving parts of the main clause to the front, but relative clauses block or prevent that kind of movement (this is sometimes referred to as the ‘subjacency principle,’ relative clauses are sometimes described as being ‘islands to extraction’ because words in a relative clause are stuck there, like Tom Hanks on a desert island). Note that a simple rule that says that the first *is* can go to the front, but not the second will work for an expression like *The little girl is leaving*; and another simple rule that says that the second *is*, but not the first, can go to the front will work for an expression like *The little girl that is talking to Susan is leaving*, but no simple rule that associates statements and questions will account for the way people actually speak. Thus, a behaviorist/associationist approach to language fails to explain readily observable language behaviors (hence, the negative agenda is accomplished).

A theory of language has to explain patterns of behavior that appear during children’s language acquisition. Children very rapidly learn the vocabulary and grammatical principles of the adult language. Children typically speak their first word at around twelve months of age, have a productive vocabulary of about fifty words when they reach eighteen months, and start using combinations of words at twenty-four months. By the age of three, children have mastered many of the fine details of their native grammar (at least in comprehension, as evidenced by the way children respond to statements such as *Push the dinosaur to the lion with the pencil* in act-out tasks, where children demonstrate their understanding by moving little toy creatures around).

The behaviorist account comes up short in the way it explains language acquisition. According to the account, language behavior should follow the same rules as other kinds of behaviors. If so, children should learn language behaviors via trial and error. In that case, children would repeat language behaviors that are rewarded and stop producing language behaviors that are not rewarded. This account predicts that feedback from care-givers would be an important factor in language development. In fact, feedback from care-givers and interaction are important contributors to language acquisition, but not in the way the behaviorist account predicts.

In observing how care-givers interact with infants and small children, child language specialists have observed that care-givers provide far more feedback for some aspects of language than others. In particular, care-givers often correct children’s mistakes when those mistakes result in semantic confusions. An example of this error would be when a small child *overextends* a word like *kitty* to mean all small furry animals. A child who calls a dog *kitty* is likely to be corrected by a competent speaker. However, small children are very rarely corrected when they produce grammatical errors. For example, the expression *I want hold Postman Pat* is ungrammatical (it is missing the infinitival marker *to*: *I want to hold Postman Pat*). This production error is likely to go either unnoticed or uncorrected in an exchange with an adult speaker. The adult may very well provide a paraphrase or an extension of the expression (*You do? Postman Pat is nice, isn’t he*?), but the adult is unlikely
to say *That’s wrong. You have to say, ‘I want TO hold Postman Pat.’* Such corrections do occur from time to time, but there is little convincing evidence that small children pay any attention to them (in that the same error is likely to appear again, even after corrective feedback is provided). Nonetheless, children’s production eventually conforms to the standard ways adults use the language. That is, children learn the adult grammar (so, among other things, children learn when to use and when not to use overt infinitival markers like *to*). This appears to be a case of learning by observation, without feedback, and without reward. (Observational learning similar to this occurs in other areas of development, including social interaction.)

These and other problems with the behaviorist view of language led to the demise of behaviorism, at least in its radical form, as a theory of language function. The cognitive revolution led to the development of theories of language that appealed to ‘hidden’ internal states and processes as the underlying causes for observable language behaviors (e.g. the kinds of sentences that people produce when speaking spontaneously; the patterns of errors that people make; the fact that some words and sentences prove harder to understand than others). These accounts appealed to a mental grammar, a set of rules or principles that governs how words can be combined in sentences and a mental lexicon, a set of words (*morphemes* might be more accurate) that are combined in production in particular ways according to the dictates of the grammar (a morpheme is the smallest unit in a language that conveys sense/meaning). This formulation, with its ‘invisible’ mental structures and processes, violates a fundamental tenet of radical behaviorism, which outlawed appeals to hypothetical mental representations and processes to explain behavior. While cognitive theories must be tied to observable events (this prevents backsliding into pure speculation and fantasy), most, if not all, contemporary cognitive scientists view the abandonment of radical behaviorism as a good move (because behaviorism does not provide a satisfactory framework for explaining the full range of observable language phenomena; and because cognitive theories, when properly constructed, are just as testable and falsifiable as any theory from the ‘dust-bowl empiricism’ tradition).

18.2.2 The second wave: experimental psycholinguistics

The second wave of psycholinguistics emerged as the result of ideas introduced by linguists (such as Noam Chomsky), ideas about human memory systems introduced by cognitive psychologists (like George Miller), and ideas about language development introduced by developmental psychologists (like Roger Brown). While linguists are chiefly concerned with descriptions of language form and structure, psycholinguists are chiefly concerned with the way humans represent language form and the processes by which they activate that knowledge in real time to communicate (to produce and understand language).

Research in psycholinguistics is organized approximately among subdisciplines that line up with important distinctions in linguistic theory. Language can be thought of as a hierarchically organized, multi-tiered system of representations and processes. Conceptually more basic units include phonemic and phonetic representations (basic units of sound in a language). Syllables, morphemes and words consist of patterns of phonemes. Phrases, clauses and sentences consist of patterns of words (or, morphemes, which may be a more appropriate description of agglutinative languages, like Turkish and Aleut). Discourses or conversations consist of sets of sentences. Pragmatic processing in language allows us to combine information about the literal semantic content of an expression with information about the speaker and situation in which that expression was uttered to determine the
Psycholinguistics

speaker’s intentions or overall ‘force’ of the message. Psycholinguists have developed cognitive theories relating to all of these levels of representation and processing. While some psycholinguists happily move between levels in this hierarchy to conduct research, others happily spend entire careers focusing on just one aspect of language.

Early theorizing in the cognitive tradition adopted ideas from transformational grammar, which was developed by linguists to do two things. First, transformational grammar represented an attempt to codify the rules that allow a speaker to combine a finite number of elements (words or morphemes) in new ways to produce an unlimited number of expressions (sentences). Second, the grammar specifies how and why different expressions can have the same meaning (I shot the sheriff and The sheriff was shot by me mean the same thing, essentially). Transformational grammar argues that the active voice I shot the sheriff and the passive voice The sheriff was shot by me mean the same thing because they have the same underlying syntactic representation (known as ‘canonical form’ or ‘deep structure’), despite their differences in surface form.

An early attempt at a psychological theory of grammar and sentence comprehension derived its predictions from transformational grammar. The derivational theory of complexity, an account of human sentence processing, proposed that some sentences differed more from their underlying forms than others, because some sentences required more mental operations (transformations) to be converted from surface form to canonical form. On the assumption that more mental operations mean more processing time, the derivational theory of complexity predicted that some sentences would take longer to understand than others. The theory also made claims about how words were organized into larger representational units (phrases, clauses, sentences), which entails claims about where the boundaries between representational units are (e.g. in the sentence The green apple fell out of the tall tree, there is a boundary between the initial noun phrase The green apple and the predicate fell out of the tall tree).

Click-detection experiments in which noises were placed either at important syntactic boundaries (between the subject and predicate, for example) or within important ‘chunks’ (e.g. between fell and out) supported some claims about the mental representation of sentences. When stimuli had clicks within important units, those clicks were often perceived as being at the boundary between two units. These results support the existence of phrases and clauses as important units in the mental representation of sentences. If sentences were perceived as unorganized sequences of words, there is no reason why clicks located at one position in reality (e.g. between fell and out) would be more likely to be perceived at another position (e.g. between apple and fell).

The derivational theory of complexity had less success in its predictions about how long processing a sentence should take. Some sentences that were more complex in theory, because they required more transformations, did not, in fact, take more time to understand (nor were they less accurately comprehended than sentences requiring fewer transformations). While ideas borrowed from linguistics – hierarchically organized mental representations for the structure of sentences, for example – proved to be useful, the specific details of specific linguistic theories like transformational grammar did not appear to succeed in predicting actual people’s behavior across a range of task environments. The take-home message for many psycholinguists at the time was that adopting assumptions straight out of linguistic theory was not particularly helpful. Others took to heart the idea that humans really did carry around a mental grammar that governed how sentences were organized in memory, and they set about trying to find a more psychologically plausible account of mental grammar.
While the details of transformational grammar, when put to the test, did not have strong appeal to many psycholinguists, other aspects of linguistic theory continued to influence psychologists working on language. For example, ideas from the linguistic study of phonology and morphology strongly influenced work on speech comprehension and word processing in psycholinguistics (and continue to do so today). In speech comprehension, theorists like Mark Liberman applied ideas from articulatory phonology to derive predictions about how people understand spoken language. Articulatory phonology makes a close connection between motor movements for speech (how the throat muscles, tongue, lips and other articulators move during speech) and the classification and categorization of speech sounds during speech perception (i.e., basic speech sounds, phonemes) can be classified according to the way different articulators move when different phonemes are produced. Psycholinguistic theories of word-recognition in speech comprehension during this era relied upon theories of phonological representation and speech production (especially co-articulation), as well as linguistically motivated theories of morphology (‘atomic’ components of complex words). The frequency ordered bin search (FOBS) model of word-recognition and lexical access argued that representations of words in long-term memory were organized around root morphemes (e.g., in the word cats, cat is the root morpheme, s is an inflectional morpheme that modifies the meaning of the root from singular, one cat, to plural, more than one cat). This account successfully predicted certain kinds of frequency effects (words with low surface frequency are recognized faster than expected if the root morpheme has relatively high frequency). It also successfully predicted differences in processing difficulty between words with real suffixes, like planter (plant + er), and words with fake or pseudo-suffixes, like sister (while a planter is someone who plants, a sister is not someone who sists).

The ‘cohort’ model is another theory of word-recognition and lexical access that appeals to linguistically motivated accounts of phonology and morphology. The cohort model proposes that word representations in long-term memory are organized according to patterns of sound. It also makes claims about how information from long-term memory behaves as spoken words are encountered.

According to the cohort model, long-term memory representations for words are organized according to their initial sounds. Each group of words is a cohort. When the initial sound in a word is identified, all of the word representations in long-term memory that match that initial sound are activated. As more sounds in the stimulus are identified, the cohort is reduced, until finally, only one matching candidate from long-term memory remains activated—typically this ‘sole survivor’ is the entry in long-term memory that most closely matches the perceptual features of the stimulus.

The cohort model predicts a number of different effects that are consistent with the available experimental evidence. For example, the partial word cap... should activate the matching entries from long-term memory captain and captive. If the word captain is activated (i.e., it is resident in short-term or working memory), that activated representation should affect the processing of stimuli that are encountered while the representation is active. If this is the case, then hearing the partial word cap... should speed up the processing of the word ship (because ship is semantically related to the activated representation of captain). For the same reasons, hearing cap... should facilitate (speed up) processing of the word guard (because captive is semantically associated with guard). In fact, when people are asked to respond to words like ship and guard, by either reading the word out loud or indicating whether it is a real word in English, they can perform those tasks faster if they have just heard the word-fragment cap... Thus, consistent with the model’s claim, partial
word information appears to rapidly and automatically activate an entire range of stored word representations (a cohort). This cohort gets weeded down as more information from the stimulus becomes available. So, if the stimulus were *capta…*, which is compatible with *captain* but not *captive*, we would expect to see priming (speeded processing) of the word *ship* but not *guard*, and this is, in fact, what happens in experiments using those kinds of stimuli. The cohort model accounts for other speech perception phenomena as well, such as the fact that words can be recognized before the end, so long as partial word information reduces the cohort to one sole survivor before the end of the word (i.e. the recognition point is somewhere in the middle of the word). Words with recognition points closer to the beginning are, all other things being equal, recognized faster than words that have later recognition points.

18.2.3 The third wave: the rise of the machines

Technological advances have had a transformative effect on psycholinguistics since about 1990. Advances in computing technology have created new opportunities for research for a wider range of researchers. Researchers have always been able to do high quality psycholinguistic research using only paper and pencil, but personal computing has made more sophisticated stimulus presentation and data collection available at very low cost. Beyond facilitating data collection, advances in computing technology have changed the way researchers analyze their data and model psychological processes. Advances in neurophysiological and neuro-imaging methods have also contributed to new understandings about how the brain responds when people produce and understand language.

Computing technology has had a profound effect on the way psycholinguists develop and evaluate theories of language function. The change is seen in the move from qualitative accounts of language function (e.g. qualitative claims; box-and-arrow diagrams) to implemented computational systems (computer programs designed to simulate some psychological process). The computational modeling tradition in psycholinguistics started with flow-chart style models, but quickly evolved into other types of computational systems patterned after real neural networks (i.e. neural network models, aka parallel distributed processing models). These mathematical models have profoundly changed the way psycholinguists view a variety of language functions, including speech perception, vocabulary acquisition and word processing (lexical access), normal and disordered reading (dyslexia), and sentence processing.

Vigorous debates continue regarding the best way to characterize some language processes. Competition between formal models in the area of reading provides a case in point that helps illustrate differences between different modeling approaches. When people read, they use visual input to retrieve phonological (sound) and semantic (meaning) information that is stored in long-term memory. The intermediate steps that occur between registering a visual pattern and retrieving the meaning that goes with that pattern are the focus of different types of formal computational models. For instance, Max Coltheart and colleague’s (2001) ‘dual-route’ model of reading proposes that readers have available to them two different means of accessing meaning from print. The first route involves the application of letter-to-sound correspondence rules (grapheme-to-phoneme correspondence, or GPC, rules). This is the mechanism that you can use to pronounce words that you have not previously encountered, like *brenk*. This mechanism will not succeed for words that are pronounced differently than they are spelled (e.g. *yacht*, *colonel*, *cough*, *pint*). To activate the phonological codes (pronunciations) that are associated with irregular sets of graphemes...
(letters), the dual-route model proposes a direct look-up mechanism. This direct or visual route bypasses letter-by-letter phonological analysis. To access the pronunciations of irregular words (like *yacht*), readers treat the stimulus as a holistic object (i.e. they do not break the word down into its individual component letters and retrieve the sounds associated with those individual letters).

The dual-route model correctly predicts that words will be processed differently depending upon details of the way they are spelled, and details of the way the GPC system as a whole maps letters onto sounds. For example, all other things being equal, words that have irregular spelling patterns towards the beginning take longer to read than words that have irregular spelling patterns towards the end. Dual-route theory accounts for these effects by appealing to the GPC analysis. GPC analysis fails to produce sensible output earlier for words that have irregular patterns at the beginning than towards the end. Dual-route theory also predicts that the two routes to pronunciation can succeed or fail independently. It also predicts that some individuals may have trouble accessing pronunciations using the GPC rules, while others will have trouble accessing pronunciations using the direct or visual route. Problems with the GPC route should lead to a pattern of function in which people can respond appropriately to known words, even if those words have irregular spellings, but those same people will not be able to pronounce unfamiliar words, even if those unfamiliar words have ‘normal’ spelling. A malfunctioning GPC route may also lead to very odd responses when an individual is asked to read a word out loud; i.e. *bake* might be pronounced as *cook* or *oven*. These patterns are indeed observed in some dyslexic individuals (people whose reading ability is significantly lower than predicted based on their overall intellectual functioning, age and access to instruction). This condition is sometimes referred to as *deep dyslexia*. A malfunctioning direct route in the same individual who has a functioning GPC route should lead to a different pattern of function. The GPC route will produce sensible output for regularly spelled words (e.g. *bake*, *lip*, *can*), but regularization errors for irregular words (e.g. *colonel* will not be pronounced *kernel*, but instead more like *kulonal*). Individuals with a functioning GPC route should still be able to pronounce words that they have not encountered before (and those pronunciations will be consistent with the general patterns of letter-to-sound correspondences). This is the pattern that is observed in people with surface dyslexia.

Work in the neural network tradition by researchers like Mark Seidenberg, Jeff Elman and Jay McClelland provides competing implemented models and competing qualitative explanations for a wide variety of reading-related phenomena, including deep and surface dyslexia. Neural network models of this type have a number of attractive features. One attractive feature is that the models’ behavior changes over time as they are exposed (virtually) to stimuli. In other words, the models learn. Another attractive feature is the relative simplicity of the processing architectures. Typically, models will have a set of input units, one or more sets of ‘hidden’ units, and a set of output units. The input units can represent a variety of stimulus features, such as phonetic features or visual features. Each input unit is typically connected to the full set of hidden units, with either excitatory or inhibitory links. If an input unit has an excitatory link to another processing unit, activating the input unit by presenting an appropriate stimulus will lead to excitation in the other unit. Inhibitory connections have the opposite effect. The hidden units are, in turn, connected to the output units. Models of this type can be ‘trained’ by presenting stimuli, letting activation flow through the network until a stable pattern is obtained, and then comparing the pattern of output that the system produces to the correct or ideal pattern of output. To the extent that the actual output differs from the ideal, the connection weights between the input, hidden
and output units will be adjusted (connection weights determine how robustly a given unit responds to signals coming from another unit).

Seidenberg and colleagues (1989) developed a neural network model of this type to capture a variety of phenomena that are observed in both normal reading and in dyslexia. This ‘triangle’ model has input units representing visual features, and a set of output units that represent semantic and phonological aspects of words. The model does not have any division that is interpretable as separate ‘GPC’ and ‘direct/visual’ routes. Nonetheless it is capable of emulating different patterns of reading problems that arise in people with deep and surface dyslexia. Those outcomes are achieved via neurologically plausible changes to the basic model. One type of change produces virtual ‘lesions’ in the network. These virtual lesions are created by removing some of the output units from the system. This can lead to the system making odd substitute pronunciations, as in deep dyslexia. Other changes, such as reducing the number of hidden units that mediate between inputs and outputs, can produce different patterns of behavior. These patterns closely match those that real readers produce. Changing the number of hidden processing units can cause the model to make errors for irregular but not regular words (as in surface dyslexia). One qualitative interpretation of this change in the model is to view it as reducing the processing resources that are available to the system, which reduces the system’s ability to map inputs to outputs. Mapping may still be possible, but the system as a whole may learn more slowly and make more errors.

A third critical technological revolution involves new ways of examining how the brain responds to language across a broad spectrum of tasks. One example is neurophysiological methods. These methods are based on the fact that the brain, as an electrochemical system, produces electrical currents as neurons fire. Different patterns of neural firing produce different electrical signals, which can be detected using non-invasive methods. Electrodes placed on a person’s scalp can pick up very subtle electrical currents that result from the action of networks of neurons in the brain. Different types of language stimuli produce different patterns of neural response, which in turn lead to different patterns of electrical current at the scalp. Researchers can present these stimuli and then measure event-related potentials (or ERPs). Differences between different stimuli or different task sets can be detected within a few hundred milliseconds after the critical stimulus is presented (sometimes sooner, depending upon the nature of the task and the stimuli).

ERP research has played a central role in the development of theories of word, sentence and discourse processing. For example, ERP research demonstrates that comprehenders very rapidly access and use information about speaker identity (Is the speaker an adult or a child?) and contextual information (What concepts have been introduced previously into the discourse? Which are most likely to be relevant at a given point in time?) to perform computations that identify word meanings (lexical access) and that tie referring expressions to previously introduced concepts (reference). ERP research has also played a large role in the development of sentence-processing theories, because it provides an index of how much processing load different components of different kinds of sentences impose on comprehenders.

Neuro-imaging research and its application to studying language function has also developed by leaps and bounds in the past decade. In particular, functional magnetic resonance imaging (fMRI) has come to occupy a central position in many discussions about language. The early days of fMRI research on language revolved around questions of where different linguistic representations and processes resided in the brain. Many of these studies presented stimuli of two different types (e.g. random strings of letters versus words; xjepvmt versus judgment) and looked for areas in the brain that responded more strongly to one type
of stimulus than the other. This particular contrast tends to pick out a part of the brain in the ventral occipital-temporal lobe junction in the left hemisphere, which has been named the ‘visual word form area.’ Some more recent fMRI research continues in this tradition (for example, some researchers are looking for parts of the brain that respond only to language, and not to other kinds of complex stimuli), but there is a developing shift in the field away from asking narrow questions about where representations are stored, or where computations are conducted (i.e. which specific parts of the cortex light up when a particular task is done on a particular kind of stimulus), to asking broader questions about how large-scale networks in the brain are organized and how different parts of the brain communicate with one another and coordinate their activity in the service of some task. This kind of functional connectivity analysis may also help us understand better how broader cognitive systems, executive control, attention, working and long-term memory, perception and action networks interface with language systems as expressions are produced and comprehended.

Notes
1 See, for example, Schizophrenia: Gerald, Part 1 at www.youtube.com/watch?v=gGnl8dqEoPQ.
2 Some psycholinguistic accounts of language function in schizophrenia attribute ‘loose associations’ in speech to rampant spreading activation within the neural networks that represent word-related information in long-term memory. These accounts therefore lay the blame for patterns of language function within a set of neural systems that are primarily concerned with language processing. These accounts would seem to predict that patterns of language use characterized by loose associations should be present only when lexical processing is compromised. Hence, we would not expect to see loose associations in other patients with executive control and attention problems (as in ADHD). Hence, there may be two separate problems occurring in conjunction in patients with schizophrenia: First, a general cognitive problem with executive control. Second, a problem within the language system itself which causes uncontrolled spreading activation within the lexical network.
3 Other concepts may also be related to both lion and stripes.
4 Some accounts attribute problems in SLI to more general cognitive deficits, such as slower speed of processing, as individuals with SLI respond slower on simple reaction time tasks. However, the speed of processing account may not provide a satisfactory explanation for why some aspects of language (grammatical suffixes of certain types) present greater problems than others; the account also does not provide a clear explanation for differences in the patterns of deficits across different languages (e.g. English versus Italian versus Hebrew; see Larry Leonard’s 1998 book, Children with Specific Language Impairment, MIT Press).
5 Childhood trauma is, sadly, a frequent cause of behavior problems in clinical populations (T. Traxler, personal communication). The link between childhood trauma and language disorder is largely unexplored (but see Cortes et al. 2005).
6 A senior psycholinguist, who shall be unnamed, tells a story of collecting reaction time data in the era before desktop computers. One of the jobs that this person did in graduate school was to use a candle to cover long rolls of paper tape with soot. These rolls of tape would be placed in a mechanical device that would spool out the tape at a fixed speed. When a stimulus was presented to the subject, the device would make a mark in the soot. When the subject responded by pressing a key, the device would make another mark in the soot. The researcher could determine how long it took the subject to respond by measuring the distance between the two marks on the tape.

Further reading
References


