

PSYCHOLOGY OF LANGUAGE

[1] The scope of Psycholinguistics

Psycholinguistics is part of the field of cognitive science, and is the study of how individuals comprehend, produce and acquire language. It mainly addresses two questions : What knowledge of language is needed for us to use language ? And what cognitive processes are involved in the ordinary use of language ? Psycholinguists are also interested in the social rules involved in language use, and the brain mechanisms associated with language.

- 1.a. Language processes and linguistic knowledge
- 1.b. Illustrations of language processing

[2] General psychological mechanisms

Both comprehension and production of language are performed within the constraints of our information processing system. The system consists of structural components (memory systems) along with a set of control processes that govern the flow of information within the system.

- 2.a. The information processing system.
- 2.b. Central issues in language processing : Serial vs Parallel processing, Top-Down vs Bottom-Up processing, Automatic vs Controlled processes, Modularity vs Interactivity.
- 2.c. Development of the processing system.

[3] Language comprehension

In this part, we will examine language comprehension at a number of levels of processing (for the convenience of exposition) : levels of speech processing, Words processing (internal lexicon), syntactic parsing and sentence understanding, connected discourse comprehension.

- 3.a. Perception of language : identification of isolated speech sounds and of continuous speech.
- 3.b. The internal lexicon : representation and organization of semantic knowledge (lexical access processes in comprehension).
- 3.c. Comprehension of sentences : syntactic and semantic psycholinguistics reaches to the construction of interpretations.
- 3.d. Discourse comprehension : Comprehension of discourse, Memory for discourse, Schemata and discourse processing.

[4] Language production and conversational interaction

Language production can be viewed as consisting in four major stages : conceptualizing a thought to be expressed, formulating a linguistic plan, articulating the plan, and monitoring one's speech. These various stages will be examined in this part..

- 4.a. Production of speech and language : Serial vs Parallel models of linguistic planning ; implementing linguistic plans (articulating and self monitoring)
- 4.b. Conversational interaction : The structure of conversations, Conversational skills, conversational settings.

LANGUAGE ACQUISITION

[1] Early language acquisition

Children's construction of language emerges from their understanding of communication prior to language ; their comprehension and production of gestures reveal a basic understanding of communication processes. Although children first acquire the sound system of their native language independently of meaning, they eventually merge it with communicative gestures to form productive speech. The development of one-word speech comprises two important developments : the acquisition of the lexicon and the use of single words to express larger chunks of meaning. Children's first word combinations reveal a structure that is neither an imitation of adult speech nor fully grammatical by adult standards.

- 1.a. Prelinguistic communication
- 1.b. Early phonology : normal and disordered development of speech perception and speech production.
- 1.c. Early lexical development : normal and disordered development
- 1.d. Early grammar : normal and disordered syntactic growth ; emergence of grammatical categories.

[2] Later language acquisition

Children's grammatical development in the late preschool years includes the acquisition of grammatical morphemes and complex syntactic structures. They also progressively extend their repertoire of both direct and indirect speech acts, and master the thematic status of given vs new information, being more able to communicate in flexible ways. Children's skills as conversationalists and narrators grow from preschool to late school years.

- 2.a. Later grammar : normal and disordered syntactic development (acquisition of complex sentences)
- 2.b. Normal and disordered elaboration of language functions (speech acts, thematic structure processing)
- 2.c. Normal and disordered development of discourse processes in children (conversational skills, narrative skills)

[3] Processes of language acquisition

Language acquisition is studied in relation to three classes of variables : environmental factors, cognitive processes and innate linguistic mechanisms. (a) Gross environmental neglect or abuse may retard language acquisition ; yet, the precise aspects of the environment necessary for normal language growth are not clear. (b) Some evidence indicates that cognitive achievements facilitate language development. (c) Children given only impoverished linguistic input are able to create communicative systems that are similar to early child language, which suggests some innate guidance in early language acquisition.

- 3.a. Contribution of environmental factors
- 3.b. Contribution of cognitive processes
- 3.c. Contribution of innate mechanisms

Ouvrages de référence :

- CARON, Jean (1989). Précis de psycholinguistique. Paris, PUF Presses universitaires de France.
- CARROLL, David. W. (1999). Psychology of Language (3rd edition). Brooks/Cole Publishing Company, ITP, London.
- CLARK, Herbert H. & CLARK, Eve V. (1977). Psychology and Language : An introduction to psycholinguistics. Harcourt Brace Jovanovich Publishers, San Diego.
- KARMILOFF, K & KARMILOFF-SMITH, A. (2001) Pathways to language : From foetus to adolescent. Cambridge, Harvard University Press. [Trad. Française : Comment les enfants entrent dans le langage, Paris, Editions Retz 2003.

Psychology of language / Psycholinguistics

Three goals in preparing these lectures:

- (a) To define basic notions that will be used to analyze both normal and impaired, oral and written language.
- (b) To present the principles of psycholinguistics in an accessible manner (taking into account recent developments of cognitive psychology and neuropsychology).
- (c) To discuss some fundamental psycholinguistic issues in reference to basic psychological mechanisms that constitute our information processing system.

Lecture 1

What is Psycholinguistics?

Psycholinguistics is the study of how individuals use (comprehend - produce) and acquire language.

Psycholinguistics is part of the field of cognitive sciences. Cognitive sciences reflect the insights of psychology and linguistics, but also of other fields as artificial intelligence, neuroscience (neuropsychology), and philosophy.

Psycholinguistics stresses the knowledge of language and the cognitive processes (or psychological mechanisms) involved in ordinary language use.

Psycholinguistics is also interested in the social rules involved in language use (// with sociolinguistics) and the brain mechanisms associated with language (// neurolinguistics).

Comments:

- (1a) Although language is essentially a social phenomenon, psychology (and therefore psycholinguistics) is principally the study of individuals. Three sets of "individual" mental processes are dealt with:

+ How we perceive and understand speech (and written language)?

- + How we construct an utterance, from idea to completed sentence?
- + How children acquire language?

(2a) As the name implies, Psycholinguistics is principally an integration of the fields of psychology and linguistics (see below), although like most interdisciplinary fields, Psycholinguistics has a rich heritage that includes contributions from diverse intellectual traditions giving rise to some controversies in how to best think of or study language and language processes.

(3a) Psycholinguistics essentially addresses two questions:

- + What knowledge of language is needed for us to use language?
- + What cognitive processes are involved in the ordinary use of language?

In a sense, we must know a language in order to use it, but normally we are not fully aware of this knowledge. We must have a tacit knowledge.

Tacit knowledge refers to the knowledge of how to perform various acts, whereas explicit knowledge refers to the knowledge of the processes or mechanisms used in these acts. We sometimes know how to do something without knowing how we do it (e.g., a tennisman might know how to serve a ball 150 kms/hour, but might have little or no explicit knowledge of the muscles that involved in this act.

Similarly, we may distinguish between knowing how to speak and knowing what processes are involved in producing or understanding speech; generally, much of our linguistic knowledge is tacit rather than explicit.

+ The knowledge question:

What knowledge of language is needed for us to use language?

Four broad areas of language knowledge may be distinguished.

Phonology:

deals with the system of sounds used in a particular language.

Syntax:

deals with the grammatical arrangement of words within sentences.

Semantics:

deals with the meanings of sentences and words.

Pragmatics:

deals with the social rules involved in language use.

+ The processes question:

What cognitive processes are involved in the ordinary use of language?

By ordinary use of language, I mean such things as understanding a lecture, reading a book, writing a letter, holding a conversation, etc. By cognitive processes, I mean processes such as perception, memory, thinking or reasoning [kinds of information processing].

The interplay of linguistic knowledge and language processes is a continuing theme of Psycholinguistics. Let us briefly examine three examples of this interplay, to illustrate both the concept of "linguistic knowledge" and the concept of "cognitive processing" in a way that highlights the types of questions a psycholinguist is likely to ask.

Example 1.

Understanding indirect requests.

What happens when we comprehend a sentence like (1):

- [1] Can you shut the door?
- [2] Will you shut the door?
- [3] It is getting cold in here.
See also
- [4] Why not shut the door?
- [5] Must you shut the door?

Literally, this sentence asks if we have the ability to shut the door, but everybody assumes the speaker is asking us (in an indirect manner) to shut the door. An interesting fact about indirect speech acts is that although there is no direct relationship between the form of the sentence and its intended meaning, listeners apparently have little trouble comprehending these speech acts.

In (1), the speaker questions the ability of the interlocutor to perform the action. Other ways could be used: for example, to question the interlocutor's willingness to perform the desired action (as in 2), or to indicate the reason why such an action needs to be done (as in 3). In (4) the speaker wants the door to be shut, whilst in (5) the speaker normally wants the door to remain open.

Why is the request phrased indirectly?

Part of the reason is that *we know certain rules* about the use of language in social settings (because we have learned them), including rules of politeness. A request is, by definition, an attempt to change another person's behaviour. This can be perceived as intrusive at times, so we soften it with indirect speech. An indirect request is more polite than a direct command such as "Pass the salt". We know this, as it is a part of our pragmatic knowledge of our language.

From a *processing standpoint*, a speaker takes this pragmatic knowledge into account when producing a sentence like [1] in a social situation. That is, the speaker utters the sentence with the understanding that it will be taken as a request. The listener

presumably shares this aspect of pragmatic knowledge and interprets the sentence as a request rather than in a literal manner.

Yet, the exact processes by which the listener arrives at the nonliteral meaning are not fully clear:

Some authors assume that the listener first extracts the literal meaning of the sentence, then decides, on the basis of the context, that the literal meaning is not what the speaker intended, before computing an indirect meaning (voir Carroll, p.227 et 139).

Other authors consider that it is not always necessary to compute a direct meaning prior to indirect meaning.; first computing the literal meaning would only be a special strategy only used when an indirect speech act is presented either in a way that does not fit with previous discourse or with no context at all.

Example 2: Sentence parsing strategies.

What happens when we understand a sentence like (1):`

[1] The teacher bores the students

Definition of parsing :

A first step in the process of understanding a sentence is to assign the elements of this sentence to linguistic categories (a procedure known as "parsing"). Figure x below (indicateur à branches, phrase marker) depicts some of the successive points in parsing a sentence like "*Le professeur endort les étudiants*". We recognize "Le" as a determiner which signals the beginning of a noun phrase; our knowledge of noun phrases is that they take the form NP → Det + (Adj) + N , so at this point we are looking either for an optional adjective or a noun. We recognize the next word, "professor" as a noun, and add it to the noun phrase, etc. (Figure x shows how the remaining items are added).

Parsing procedure:

We may think of parsing as a form of problem solving or decision making, in the sense that we are making decisions (not necessarily in a conscious manner) about where to place incoming words into the phrase marker we are building. Various principles can be taken into account to make such decisions; for example:

- + Immediacy principle (= we make the decisions immediately as we encounter a word)
- + Wait and see principle (= wait for further information before deciding between several possible interpretations).

Parsing strategies:

If we are making decisions about where words fit into the syntactic structure of a sentence, what are these decisions based upon? The answer is: we use strategies.

Late closure strategy.

This strategy states that, wherever possible, we prefer to attach new items to the current constituent. For example, in "Tom said that Bill had taken the cleaning out yesterday", the adverb "yesterday" may be attached to the main clause (Tom said yesterday) or to the subsequent subordinate clause (Bill had taken the cleaning yesterday). Findings show that we prefer this latter strategy.

Further evidence for the late closure strategy comes from eye fixations of subjects reading structurally ambiguous sentences such as "Since John always jogs a mile seems very like a very short distance to him". The subjects' eye fixation times on the last few words were much longer than on the earlier ones, implying that readers had misinterpreted the phrase "a mile" and had to make some later adjustments. In such "garden path sentences" we interpret a sentence in a particular way, before finding out near the end that we were wrong.

Minimal attachment strategy.

This strategy states that we prefer attach new items into the phrase marker being constructed using the fewest syntactic nodes consistent with the the rules of the language. For example, a sentence fragment such as "Ernie kissed Marcie and her sister ..." could be interpreted as either a noun phrase conjunction (Marcie and her sister were recipients of a kiss), or the beginning of a new noun phrase (and her sister gave her a present). According to the minimal attachment, we prefer the former interpretation. Similarly, reading times were faster for (1) than for (2), (1) being consistent with minimal attachment, whilst (2) requires building a new constituent (see figure 6-2, page 132, Carroll).

- [1] The city council argued the mayor's position forcefully
- [2] The city council argued the mayor's position was correct

The parsing mechanism we have considered has some clear characteristics (that have been debated extensively in recent years).

- [1] For one thing, it is assumed to be based on structural preferences which are themselves presumably based on information processing limitations.
- [2] For another, it can be described in modular terms; that is it does not directly interact with other parts of the language comprehension system.

Cognitive processes involved in parsing procedure:
Current debates

Structural or lexical preferences ?

The assumption in the above description of the parsing has been that the parsing mechanism is based on structural preferences that guide the parsing procedure.

An alternative approach is to give greater emphasis to lexical preferences. In this view, the linguistic structure of individual lexical items includes knowledge of the kinds of syntactic arrangements in which they may participate.

Some authors argue that a sentence is initially analyzed in accordance with the strongest or preferred lexical form of the verb. For example, although sentences (1) and (2) appear to be similar, the interpretation favored in (1) is that of a simple noun phrase (the dress), whereas the verb "wanted" tends to favor a more complex noun phrase (the dress on the rack).

[1] The woman positioned [the dress] on the rack

[2] The woman wanted [the dress on the rack]

See other examples in Carroll p. 133

Conclusion: whether lexical factors operate in initial parsing or at a subsequent stage (for deciding whether a previous parsing is correct) is still a matter of much debate.

Modularity or interactive parsing?

The above description of the parsing is consistent with a modular approach to language comprehension in which comprehension as a whole is the result of many different modules, each devoted to a particular aspect or treatment.

In a modular view, parsing is performed initially by a syntactic module that is not influenced by any other variables (such as the meaning of the sentence, the subjects' general world knowledge, etc.).

An alternative view is that syntax and semantics interact during the comprehension process. In a study by Tyler & Marslen-Wilson, for example, subjects heard sentence fragments containing structurally ambiguous phrases, with the first clause of each fragment biasing the listener to one of the readings; for examples, (1) et (2) were two such fragments:

[1] If you walk too near the runway, landing planes . . .

[2] If you have been trained as a pilot, landing planes . . .

At the end of the fragments, a word was presented visually that either was or was not an appropriate continuation of the sentence: "are" is appropriate for the first fragment, and "is" for the second one. The time the subjects took to name the words was shorter for appropriate words than for inappropriate ones, suggesting that the subjects were able to use the context to assign one or another structural meaning to the sentence immediately.

Yet, several other studies (Rayner et coll; voir Carroll, p. 135) have supported a modular view according to which the parsing mechanism is viewed as a syntactic module that operates on specific input (a string of words) to produce a specific out put (a tree structure); in this view, the role of context is not directly related to the initial parsing of a sentence, but rather to its subsequent interpretation.

Example 3: Impaired language

The primary focus of Psycholinguistics clearly is on language processes in normal individuals. However, we can learn a great deal about language by studying individuals (children or adults) with impaired language functioning.

An Aphasia for instance is a language disorder due to brain damage. One type of aphasia, called Wernicke's aphasia, involves a breakdown in semantics. For example, consider excerpt (1) below:

[1] Before I was in the one here, I was over in the other one.
My sister had the department in the other one. (Carroll, 10)

What can we learn from this excerpt? The semantic relationships between words in this excerpt are seriously disrupted, suggesting that the patient's semantic knowledge has been impaired by the brain damage. In contrast, phonological knowledge was spared; the speech, although devoid of meaning, was articulated smoothly, with appropriate pausing and intonation. It also displays appropriate syntactic structure.

The study of the relationship between the brain and language is called neurolinguistics. Much is still to be learned about this relationship, but what is presently known is instructive. Depending of the exact location of the injury, and of its severity, an individual who has suffered from a brain injury may display a wide variety of reactions. One individual may have normal comprehension but be deficient in language production. Another individual may have no loss of ability with sentence structure but have greater than normal problems finding words. Still other individuals may be unimpaired in comprehension and production but be unable to repeat exactly what they have heard and understood.

In normal individuals with intact brains, various facets of language -sentence structure, meaning, sounds- appear to form a smoothly coordinated system of communication; in brain-damaged individuals, however, this system is revealed to be a combination of separate

parts (for the deficits in such persons are nearly always selective rather than total).

Thus, brain injuries enable us to analyze an apparently unified program of language abilities into its separate components and raise question about how such abilities become integrated in the normal adult.

Example 4:

Language in children

Studying language acquisition, both in normal and troubled children, can also improve our knowledge about cognitive processes involved in ordinary use of language. This will be addressed in much detail in several lectures specifically devoted to early and later language acquisition.

Lecture 2:

Psychological mechanisms and language processing

Psycholinguistics focuses on how people process language -how we comprehend and produce spoken and written language- and how these skills are acquired. In order to understand these language processes, we need to understand the major properties of language as well as the processing characteristics of the individuals who use it.

For this lecture, basic notions about the structure of language as well as basic grammatical concepts will be assumed already known.

The following concepts will be taken for granted.

Linguistic productivity: this notion refers to the fact that there is no limit to the number of sentences in a language.

Duality of patterning: words are composed of phonemes which, in turn, are composed of distinctive features. In each instance, the smaller units are combined in a rule-governed manner to produce the larger units.

Morphology: words consist of one or more units of meaning (morphemes). The system of grammatical morphemes in a language provides speakers with a way of signaling subtle differences in meaning.

Phrase structure: phrase structure rules codify our intuitions about the groupings of words in a sentence. Some sentences are ambiguous

Linguistic investigations place a strong emphasis on linguistic structures of various sorts, such as phrase structure, distinctive feature, morphological structure, etc. These investigations reflect the belief that a fuller understanding of human language will reveal deep insights into the human mind. However, linguistic investigations typically focused on what we have called the knowledge question: what kinds of knowledge underlie language use? As we have seen in Lecture 1, this is different from the process question of how this knowledge is utilized. Given a particular phrase structure for example, there are still any number of ways we might comprehend or produce a sentence with that structure. More specifically, some of these ways might be preferred over others for cognitive reasons: they might be easier, or pose less burden on memory, and so on. It is therefore indispensable to examine the psychological mechanisms that are involved in language use.

In this lecture, we shall first present the general cognitive architecture of the human information processing system. Afterwards, we will examine several ways in which linguistic information can be handled by this system.

In short:

Why to study both the structural and the functional aspects of the cognitive system?

[1] Comprehending and producing language are performed within the constraints of our information processing system. This system consists of three structural components - sensory memory, working memory, and permanent memory- along with a set of control processes that govern the flow of information within the system.

[2] As concerns language processing, number of issues have been raised and should be examined in details:

- 2.a. Whether we primarily use serial or parallel processes
- 2.b. Whether we use Top-Down or Bottom-Up processes
- 2.c. Whether language processes are automatic or controlled
- 2.d. Whether language processing is modular or interactive

[3] Children appear to process information very differently than adults, but studies of the development of the processing system suggest that most of the system is developmentally invariant.

Next pages will be devoted to a survey of these fundamental issues. The following three major points will be examined:

- A. The information processing system
- B. Central issues in language processing
- C. Development of the processing system

A. The information processing system

Figure z below presents a general model of information processing that consists of three mental structures and a set of processes that move information from one structure to another.

Insert Figure z here

According to this model of mental functioning, environmental information entering into the cognitive system is successively encoded, stored and retrieved by a set of distinct mental structures. The emphasis is on the flow of information through the system. Let's first briefly give an overview of this model, before examining its relevance regarding language processing.

Sensory stores

The sensory stores take in the variety of sensory events to which we are constantly exposed (colors, tones, tastes, smells, etc.), and retain them for a brief period of time in a raw, unanalyzed form.

Sperling's princeps studies.

The visual sensory store has been studied by Sperling (1960). He presented participants with arrays of letters or digits (four columns and three rows) for very brief periods of time (50 milliseconds). Immediately after the presentation, participants were asked to report what they saw. Typically, they could report only 4 or 5 of the 12 letters. However, many participants reported that they could see all the letters in the array but lost some of them by the time they reported the others.

This finding provided a challenge for Sperling, for if participants did indeed retain all of the items for such a brief period of time, then the procedure of asking the participants to report all the items was likely to underestimate their retention. Sperling's response to this problem was ingenious. He devised a *partial report technique* in which the participants had to report only a portion of the entire array. Participants were presented with a high-pitched tone if they were to recall the top row of the array, a medium-pitched tone for the middle row, and a low-pitched tone for the bottom row. The tone was presented immediately after the array was removed. Sperling's reasoning was that if participants did not know which row they were to report, then they would have to retain the entire array for a brief time. The results were that the participants reported about 3/4 of the row under this partial report condition, suggesting that they were able to retain about three-fourths of the entire array. Sperling concluded that a considerable amount of information in a stimulus array is available immediately after it is presented.

Further studies examined the duration of storage in the visual sensory store. This was done by delaying the presentation of the tone signal. When the tone was delayed for only 1 second, the performance in the partial report condition declined to 36%. Sperling inferred that information in the visual store persisted for approximately 1 second.

Comparable studies of the auditory sensory store (Darwin et al. , 1972) showed that the advantage of partial report over whole report persists for approximately 4 seconds. This difference might well be important in language processing. When we are processing language visually (typically in reading), we can always go back and re-read, but this not possible with speech. The relatively longer duration of the auditory store may enable us to re-analyze auditory messages that were not understood initially.

Working memory

The second type of memory has been traditionally referred to as short-term memory (STM), and more recently as working memory (WM). Although the meaning of the words are similar, there is a subtle difference between them.

The need for short-term storage is easy to state. The contents of the sensory stores are held for at most a few seconds, but many cognitive acts require that we hold onto information for longer periods of time. Many apparently simple acts, such as solving problems in your head, mentally retracing a path after losing your keys, remembering the topic of a conversation after a distraction, etc. are in fact a complex series of decisions, and it is necessary to have a temporary holding place for intermediate decisions.

Short-term memory is severely limited in size, or capacity; it can hold approximately 7 plus or minus 2 items of information. As a consequence, we cannot take in a string of digits such as 065848345071591918 and repeat it verbatim unless we reduce the strain on short-term memory. One way to do this is by **chunking**, in which we group individual pieces of information into larger units. With a little practice, nearly everyone can learn to chunk a 18-digit number into units such as 065-848-345 and 071-591-918. This arrangement is much easier to remember (at least for me, since these are the phone numbers of two of my children).

Working memory differs from short-term memory in that the term WM conveys a more dynamic view of memory processes. STM was usually viewed as a passive repository of information; WM has both storage and processing functions.

- + **The storage function** is similar to the storage credited to short term memory: we hold on a limited amount of information for a limited amount of time.
- + **The processing function** is related to the concept of processing capacity. Processing capacity refers to the total amount of cognitive resources we may devote to a task (and this amount is assumed to be limited).

Working Memory

Example 1.

Suppose you are asked to multiply 8 times 4 in your head. This would pose no problem. But if the numbers were 84 and 67, the task would be much more difficult. Part of your processing capacity would be devoted to performing the arithmetic operations of multiplying 4 by 7, and part of this capacity would be needed to retain the result (28) in temporary storage while you multiply 8 by 6. Ultimately, these two functions would be in conflict with one another, competing for the same limited pool of resources.

Working Memory

Example 2.

Alpha-span: words are presented in random order but must be recalled in alphabetical order.

Permanent memory.

Permanent memory, which is also known as long-term memory (LTM), is a repository of our knowledge of the world. This includes general knowledge (including the rules of grammar or of arithmetic, along with personal experiences such as memories of our childhood and adolescence).

Tulving has distinguished between two types of permanent memory: semantic memory and episodic memory.

+ **Semantic memory:**

Refers to our organized knowledge of words, concepts, symbols and objects; it includes our general knowledge (grammar, arithmetic), spatial knowledge (the typical layout of a house), social knowledge (how and when to be polite), etc. Semantic memory holds the information that is not tagged for a particular time or place (e.g., it holds the information that horses have four legs and a tail but not the memory of the last time you went horseback riding)

+ **Episodic memory:**

Holds traces of events that are specific to a particular time and space. This is the memory we use to keep a record of our personal experiences; it includes such items as what you had for breakfast this morning, what you were doing when you learned a man walked on the moon, or where you got your first job. (As these examples illustrate, episodic memory varies from person to person, and is constantly updated)

+ **Procedural memory:**

Is sometimes distinguished from the other two memory structures. It specifically includes information about motor skills (typing, swimming, bicycling, etc.)

Relevance of these mental structures for language processing

How do these concepts (sensory stores, short-term and working memory, long-term memory) apply to the problem of how we comprehend spoken language?

In comprehension, we may assume that as we hear a sentence, the sounds are first stored very briefly in the auditory **sensory store**; the sounds are held in this store for about 2 to 4 seconds, which gives us time to recognize an auditory pattern, i.e. :

Recognize speech sounds (= identifying acoustic cues that are present in the speech signal)

Organize the sounds into syllables and words (but it is not clear when and how that happens, see below Bottom-Up vs Top-Down processes)

Working memory has two functions: storing and processing. As regards its *storing function*, WM is only able to hold about seven units of information. Since many sentences are longer than seven words, we need some way to deal immediately with more than seven words. One way we do this is to chunk the words into grammatical constituents such as Noun phrase [e.g., My sister, My sister's boy, My sister's little boy, etc.] or Verb phrase [e.g., Bought a book, Bought a book of deserts, Bought a book of chocolate deserts, etc.], thereby reducing the storage burden to perhaps two or three constituents.

The *processing function* of WM is used to organize the words into constituents, and to build the phrase structure of an incoming sentence. As new information enters working memory, some of the older information is thus reorganized into larger units, other information is lost, and still other information is sent to permanent memory (where the resulting memory trace has both episodic and semantic attributes)

Permanent memory plays several roles.

[a] *Semantic* memory contains information on the words that we retrieve during pattern recognition.

[b] While this process is going on, we are also building up an *episodic* memory representation of the ongoing discourse. That is, once we complete the processing of a given sentence, we extract the gist of it and store it in episodic memory

Conclusion

Mental structures and processes that allow the human mind to encode, store and retrieve information can be described independently of language. Modelisation of these structures and processes, however, provides a framework for understanding how language processing occurs.

Although it is generally agreed that we encode, store and retrieve linguistic information along the general lines sketched above, the specific processes have yet to be addressed.

B. Central issues in language processing

In this section, we examine several alternative ways in which linguistic information can be handled by the information processing system that we have just sketched above. Different types of process will first be presented and discussed individually, before being applied to an extended example of language processing.

Serial and parallel processing

Serial processing refers to processes that take place one at a time. *Parallel processing* refers to processes two or more of which take place simultaneously.

Suppose we wish to develop a model of language production. The starting point is the idea that the speaker wants to convey; the ending point is the actual articulation of the idea. But what happens in between?

- + *A serial model* would divide the process into stages. There might be a stage devoted to building up the phrase structure of the sentence, another stage devoted to retrieving the lexical items that are inserted into that structure, and still another stage devoted to determining the correct pronunciation of these lexical items. The serial model would assume that these stages occur one at a time, with none overlapping.

- + *A parallel model*, on the contrary, would assume that all of these processes could take place at the same time. That is, we could be phonetically specifying one word while we search for the next word, or both of these processes could take place as we flesh out the syntactic structure.

Language examples

Take the examples shown below (Carroll, Fig. 3-2, p. 55)

In example [1], we interpret the middle letter as an /h/ in one word, but as a /a/ in the other, despite the fact that the letter is physically identical in the two cases.

The other examples [2] to [5] show degraded letters. It is not difficult, however, to identify what the word is in each case. The three degraded letters in example [2] can be identified respectively as /r/, /e/ and /d/; but the same three degraded letters appearing in the other examples are rather identified as /p/, /f/ and /b/.

At first glance, all this may appear to be paradoxical. It seems reasonable to say that we are using the context to help decide the identity of the degraded letters. However, that context is a word, and we normally think of first identifying the letters before identifying the word. How can we use the word to help identify the letter?

The answer lies in parallel processing. Assume that we are identifying the individual letters, and at the same time, actively trying to fit the letters into various possible words. Some of the identified letters enable us to recognize the word as a familiar word, and then we identify the obscured letter from our knowledge of the spelling of the word. That is, we are processing at the letter and word levels simultaneously.

Serial models have been influential in the study of cognition from 1960 to 1980, in part because many of the models were based on an analogy with the electronic computer which tends to execute processes rapidly in a serial manner.

Parallel distributed processing models have been described as neurally inspired because they use the brain, rather than the computer, as the dominant metaphor. Some authors (Rumelhart & McClelland, 1986) theorized a cognitive model as a vast, interconnected network of information nodes with each node influencing and being influenced by a large number of adjacent nodes.

Top-Down and Bottom-Up processing

Suppose you are listening to a lecturer, trying to comprehend what s/he says and remember the main points of the lecture. Your language processing can be analyzed as occurring at several levels.

At the lowest level (the phonological level), you are identifying the phonemes and syllables that the teacher is using.

At a higher level (the lexical level), you are using those phonemes and syllables to retrieve the lexical entries of the words from your semantic memory.

At the next level (the syntactic level), you are organizing the words into constituents and you are forming a phrase structure for each incoming sentence.

At the highest level (the discourse level), you are linking the meaning of a given sentence with preceding ones, and therefore you are organizing sentences into higher-order units.

Bottom-up processing is a processing which proceeds from the lowest level to the highest level of processing in such a way that all of the lower levels of processing operate without influence from the higher levels. [This means that the identification of phonemes is not affected by the lexical, syntactic or discourse levels; it means that the retrieval of words is not affected by syntactic, and so on].

However, as we have already seen about the serial and parallel processing, there is some reason to doubt that a strict bottom-up model will provide a fully comprehensive account of how we understand language.

A Top-down processing model, in contrast, states that some information at the higher levels may influence processing at the lower levels. For example, a sentence context may influence the identification of words within that sentence. Speaking more intuitively, a top-down model of processing is one in which one's expectations play a significant role.

N.B. There is some relationship between the preceding two alternative types of information processes. A Top-Down process is often (but not necessarily) a parallel process, and a Bottom-Up process is usually serial.

Automatic and Controlled processes

When we discussed the role of WM working memory along within other mental memory structures, we introduced the idea that we may have a fixed processing capacity for handling information. This notion has been a central assumption in a variety of accounts of human cognitive functioning. It is an important concept when considering human performance on complex tasks, such as language processing. When the task is complex, one part of the task may draw substantial resources from this limited pool of resources, thereby leaving insufficient resources for other parts of the task, resulting in overall impaired performance.

Tasks that draw substantially from this limited pool of resources are called controlled tasks, and the processes involved in these tasks are referred to as "controlled processes".

Tasks that do not require substantial resources are called automatic tasks, and processes that do not require extensive capacity are referred to as "automatic processes"

Various criteria have been used to determine whether a task is automatic or controlled. These criteria include sensitivity to developmental and strategy effects. For both these criteria, controlled tasks are more sensitive than automatic tasks; that is, automatic tasks appear to be unrelated to the age of the individual or to the strategy employed. Several tasks, however, become automatic as a consequence of our degree of practice with them (example: tying our shoes): they were more demanding when we were young and have become automatic through practice.

Illustrations. Different examples will be used to illustrate this distinction, in various domains of spoken and written language. (E.g., one language processing task that is automatic, at least for adults, is recognizing common words, most probably due to our large amount of experience with words. In contrast, building up a phrase structure for a sentence is a controlled process.)

Modularity of the language processing system

Within linguistics, the concept of modularity refers to the independence of the different linguistic subsystems within the grammar. Within cognitive psychology, modularity refers to the degree of independence of the language processing system, taken as a whole, from the general cognitive system.

The modularity position is that the language processing system is a unique set of cognitive abilities that cannot be reduced to general principles of cognition.

The alternative position stresses the interconnections between language and cognitive processes by emphasizing the role of concepts such as working memory, automatic processing, and parallel processing in language comprehension, production and acquisition.

The notion that language is modular is related (but not identical) to the argument that our "language faculty" is biologically innate. Certainly one way to talk about "modules" is to talk about innate modules, but this is not a necessary property. A "module" is dedicated to performing one aspect of a complex task: whether this assignment is biologically given or acquired through experience is a separate issue.

An example of language processing

Up to now, we have briefly presented some distinctions that are relevant to language processing. Let us now examine a specific example and see how these distinctions might apply.

[1] *I was afraid of Ali's powerful punch, especially since it had already laid out many thougher man who had bragged they could handle that much alcohol. (see Clark & Clark, 1977, p. 81)*

Sentence [1] is another example of a "garden path sentence" (see above in the Parsing strategies section). The key word here is "punch", which can mean either an alcoholic beverage or a boxing punch. The subjective impression for most people at the end of the sentence is: (a) I have assumed the wrong meaning (most probably because of "Ali" who is a famous boxer), and (b) I have to backtrack. A processing model that would account for such impressions might look like this: when we encounter a word that has more than one meaning, we survey the immediate environment of that word, make a rapid decision as to the most appropriate meaning, and then stay with that meaning unless it becomes obvious that we are in error.

This model corresponds reasonably well with subjective impressions, but are these impressions accurate? This model assumes serial processing (one meaning at a time), with top-down processing playing only a limited role (decision is based on immediate context, not the entire sentence). Because the emphasis is on decisions the comprehender must make during the course of comprehension, the model emphasizes controlled processes more than automatic processes. Finally, this approach can be safely described as nonmodular; it relies on our general ability to figure things out, not on a specialized ability that is related to language; it might even be described as common sense.

We could, however, develop a completely contrasting model. We could begin with the assumption that people routinely and simultaneously activate more than one meaning of an ambiguous word from semantic memory. We could also further assume that the retrieval of multiple meanings is a fixed property of the lexicon, that

it is automatic, modular, and bottom-up (i.e., not related at all to the sentence context).

Although this latter model may sound counterintuitive, there is some psychological evidence in favor of it. It appears indeed that we automatically activate all the meanings of an ambiguous word at least briefly; but it also appears that we decide among the choices rather quickly, perhaps within three or four words. Thus, there may be two stages of processing: an automatic stage in which all meanings are retrieved and a more controlled stage that is more top-down in nature.

The notion that we might have two different ways of approaching a sentence with an ambiguous word is not limited to this one particular example. This state of affairs is the rule in human information processing, in which we nearly always have multiple ways of doing things, and in which we usually employ the easiest, fastest, or most efficient strategy that will work. All this shows that if we are to develop a solid knowledge of how language processing takes place, we will need to rely not on introspection or subjective impressions, but rather on systematic experimentation.

C. Development of the processing system

In order to understand language acquisition, it would be helpful to understand the cognitive abilities children bring to the task of acquiring their native language. To this aim, let us ask to what extent the information processing system that we have outlined above is operating during the first few years of life.

It is clear enough that children encode, store and retrieve a great deal of linguistic information in their first few years. We therefore may begin by assuming that the information processing system (as discussed above) is, for the most part, developmentally invariant. "For the most part" only, because some components, by definition, must change; the episodic portion of the permanent memory is a repository of one's experience, and thus will clearly grow with experience. But, on the whole, we may tentatively assume that the sensory stores, working memory, and capacity limitations we have emphasized are present to the same extent in newborn infants as in mature adults.

This may seem an unlikely hypothesis, for we ordinarily think of children and especially infants as cognitively very different creatures than adults. Recent research, however, has suggested that there are some important cognitive similarities between children and adults.

Perceptual processing.

Although virtually no research has addressed the development of the sensory stores per se, there is a substantial literature on perceptual processing defined more broadly. Prior to the early 1960s, it was generally believed that the infants perceive the world in a disorganized way.

"The baby, assailed by eyes, ears, nose, skin, and entrails at once, feels it all as one great blooming, buzzing confusion" (W. James, 1890, p. 488).

Subsequent research, however, indicates otherwise. In a research by Fantz (1963) for example, babies were shown various forms while

observers recorded how long the infants looked at each one. The infants look at some patterns more than others. Fantz interpreted this finding as meaning that the infants could distinguish among the forms and preferred some to the others. Fantz found that young infants preferred patterns such as faces or checkerboards over plain stimuli.

More details about experimental techniques and paradigms to study infants' early visual preferences in Karmiloff-Smith & Karmiloff, 2001, p. 40-46

Just as the human face, for whatever reason, fascinates infants, so does the human voice. Various studies indicate that infants are biologically prepared to distinguish between speech and other sounds. For instance, Condon & Sander (1974) have shown that infants make synchronized movements (e.g., raising a finger, kicking a leg, etc.) to individual speech sounds shortly after birth. Frame by frame analysis of these movements indicated that they were synchronized with the syllables in the speech signal. Newborns show this synchrony with different languages but not with nonspeech sounds like the tapping of a pencil.

A famous study by Eimas et al. (1971) showed that infants make distinctions between different speech sounds very early in life. In this study, they examined infant's perception of the voicing contrast in English (contrasting /ba/ for which Voice onset time VOT = 0 ms and /pa/ for which VOT = 40 ms). Eimas and colleagues used a technique known as habituation paradigm to study infant's responses to these stimuli. (See Karmiloff-Smith & Karmiloff, 2001, p. 20 seq.).

Habituation refers to the decline in a response to a stimulus following repeated presentation of the stimulus. The response studied is referred to as nonnutritive sucking and involves a special pacifier that is connected to a computer that measures the number of sucks an infant makes. When /ba/ was initially presented, the infant showed great interest, and the sucking rate was high; but with repeated presentations, the sucking decreased. At this point, a new stimulus was presented. On some occasions it was a stimulus from the /ba/ category (VOT between 0 and 25 ms), and on other occasions it was a stimulus from the /pa/ category (VOT between

25 and 40 ms). The interesting result was that sucking returned briefly to a high level (dishabituation) if the new stimulus was from a different category, but not if the new stimulus was from the same category.

Non-nutritive sucking technique

Insert here Figure x in De Boysson-Bardies, p. 233

Taken together, studies of perceptual processing by infants reveal infants to have rather sophisticated perceptual skills at birth or shortly after birth. These results at least suggest that infants are well equipped at birth to process linguistic stimuli.

Short-term STM and Working memory WM

A second aspect of the cognitive system that has been examined developmentally is short-term memory. A common test of STM in grade school children is the simple memory span test: the children are presented with a list of numbers and asked to recall them (some times in forward order and sometimes in backward order). It is clear that there are significant differences between older and younger children on this task. In fact, the simple memory span test is a standard feature on intelligence test for that very reason.

See Dempster, F.N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, 89, 63-100.

These results, however, do not tell us much about the development of working memory. WM indeed consists of both storage and processing functions, and these compete for limited resources. Thus, a development change in either of these functions would affect the other. In particular, if children become more efficient in processing numbers, fewer resources would be needed for that function, thereby leaving more resources available for storage. In this way, simple memory span might improve without any change in the total size of working memory.

A series of elegant experiments by Case et al. (1982) has confirmed this analysis.

In their first experiment, 3- to 6-year-olds were given a storage and a processing task. The *storage task* was a simple memory task of remembering a list of seven words. The *processing task* was a repetition task, in which the children had to repeat auditory presented words. The investigators measured the time between the presentation of the word and the onset of the child's vocal response. The same words were used in both tasks. Case et al. found a negative correlation between the two tasks: as response time increased (reflecting processing load), memory span decreased. Older children remembered more words than younger children, but when their reaction time was taken into account, there were no differences between the two age groups.

A second experiment presented adults with the same two tasks, using non-sense words (with which adults have no familiarity). Adults' reaction time to non-sense words was slow, comparable to younger children's reaction time to real words. If memory span depends upon processing speed, then adults' recall of the non-sense words should also be similar to the younger children's recall of real words, and it was. The studies thus demonstrated that when adults are slowed down to children's speed of processing, memory span differences disappear.

Case and colleagues concluded that there is no substantial increase in overall working memory capacity with development, at least from 6 years to adulthood. Under this interpretation, simple memory span shows a functional increase in storage due to the greater efficiency of processing. This notion of developmental changes in processing efficiency is highly relevant to understanding early reading by children.

Conclusion

Globally speaking, observed differences between children and adults on cognitive tasks have more to do with the use of information processing resources than with qualitative shifts in thinking (as it is somehow suggested by Piaget's interpretation of cognitive development). The major change over time is an increase in efficiency, or a decrease in the amount of cognitive capacity required for a particular task. Increased efficiency enables a child to do simultaneously two or more tasks, or a complex task with two or more components. This is directly pertinent to language acquisition. Indeed, in order to acquire a language, a child must perform correct analyses simultaneously on the phonological, morphological, syntactic, semantic and pragmatic levels. A child of 3 years old may approach the task in a way that is similar to adults, but less efficient. For example, when working on new syntactic structures (which is a demanding task), the child may compensate by attending less to phonological specification. Perhaps only later when the syntax becomes a little easier, will the child be able to master both levels simultaneously.

La compréhension du langage

Remarques préliminaires.

(1) En toute rigueur, le langage est une faculté humaine, ou un système structuré qui se distingue de la parole (mise en oeuvre de ce système dans des énoncés). Ce qui est compris, ce sont des énoncés; il serait donc plus exact de parler de "compréhension d'énoncés" ou de "compréhension de la parole".

(2) Il faut distinguer "comprendre 1" appliqué à un énoncé (par ex. comprendre la phrase "Les oiseaux chantent dans le jardin"), et "comprendre 2" appliqué à un apprentissage (par ex. comprendre un cours de biologie). Dans "comprendre 2", il s'agit de l'élaboration de concepts et de propositions, et de leur intégration dans le système des connaissances d'un individu. Bien que cette activité cognitive se situe souvent dans le prolongement de la compréhension d'énoncés (comprendre les énoncés qui constituent un cours ou un chapitre de livre), elle est différente de la simple compréhension courante d'un énoncé dont nous allons parler ici, en essayant de montrer comment l'esprit/cerveau produit du sens.

(3) Pour étudier cette compréhension, on dispose d'un observable (l'énoncé) que l'on va traiter comme une des sources causales des phénomènes qu'on veut étudier.

3.a. Dans une première phase, l'énoncé présenté à un compreneur Z est saisi perceptivement (perception auditive ou visuelle). Nous passerons assez vite sur cette phase de perception de la parole.

3.b. Dans une seconde phase, l'information acquise par la perception (et qui a donc déjà reçu une première transformation: par exemple, identification des phonèmes, c'est à dire transformation du flux acoustique continu en une séquence de phonèmes) doit maintenant être traitée sémantiquement par l'esprit/cerveau du compreneur.

Pour décrire cette seconde phase, on peut parler de;

- comprendre l'énoncé
- interpréter l'énoncé
- extraire le sens de l'énoncé
- construire le sens de l'énoncé
- calculer le sens de l'énoncé

L'aboutissement final de cette seconde phase, c'est un fragment de sens dans l'esprit du compreneur Z: c'est le "sens de l'énoncé pour ce compreneur"; il s'agit de quelque chose, d'une "entité" qui n'est pas directement observable. Mais on a inventé des techniques de recherche qui donneront accès à ses propriétés principales (cfr. analogie avec les cailloux si petits qu'ils sont invisibles à l'oeil nu, mais qu'on peut étudier à partir des ronds qu'ils font si on les jette dans l'eau).

(4) Tous ceux qui s'intéressent au langage (psycholinguistes, neuro-linguistes, linguistes, philosophes, etc.) s'accordent sur l'idée générale qu'expriment les formules ci-dessus: la compréhension d'un énoncé est une construction de sens.

4.a. Nous parlerons d'une part des mots et de leurs "significations": celles-ci sont des représentations sémantiques à long terme, stockées dans la mémoire à long terme; le mode de fonctionnement propre de ces représentations est de pouvoir être activées.

4.b. Nous parlerons d'autre part de "sens" pour désigner une autre catégorie de représentations sémantiques: celles dont nous allons voir qu'elles ne peuvent exister qu'après avoir été construites.

La compréhension d'un énoncé, c'est-à-dire son traitement cognitif dans l'esprit/cerveau d'un compreneur, y produit causalement une suite de processus et d'états mentaux transitoires qui aboutissent finalement à la construction d'une représentation sémantique terminale qui est composée et structurée: c'est cette représentation qui constitue le sens (individuel) de l'énoncé. En parlant de "construction" du sens, on exprime en réalité 3 idées: causalité (cognitive), composition, et structure (sémantique).

(5) La causalité cognitive sous-jacente à la compréhension se réalise dans un traitement "incrémental". Par "incrément", il faut entendre "une petite opération d'adjonction à la construction en cours"; il s'agit d'une opération de durée très très courte (exprimable en dizaine de millisecondes).

Par exemple, au cours de la lecture de la phrase "Le chat poursuit la souris", l'interprétation du fait que "chat" se trouve avant le verbe "poursuit"

apporte un incrément qui permet de comprendre que c'est le chat qui poursuit (et non l'inverse). Dans les langues à déclinaisons, ce serait la terminaison de "chat", indicatrice du rôle de "chat", qui fournirait cet incrément.

L'idée est que l'information qui sert à la compréhension est triatée au fur et à mesure (Il s'agit de l'information qui se trouve dans l'énoncé, sous forme de mots et de données grammaticales, mais aussi celle qui appartient à la situation dans laquelle est reçu cet énoncé). Elle est d'abord saisie perceptivement, identifiée et reconnue, puis elle est interprétée de façon immédiate par petits morceaux -ce sont les incréments- et finalement assemblée sémantiquement dans une représentation sémantique d'ensemble qui constitue les "sens construit" de l'énoncé.

Cette conception dynamique de la compréhension, avec son caractère "au fur et à mesure", incrémental, est largement admise aujourd'hui. Elle s'oppose à celle qui a eu un certain succès il y a quelques années, dans le contexte des grammaires génératives transformationnelles. Selon cette autre conception (abandonnée aujourd'hui), le compreneur commençait par traiter la structure syntaxique de l'énoncé; ce n'est qu'ensuite qu'il y insérait l'information sémantique véhiculée par les mots.

(6) La compréhension (= la construction du sens) s'effectue à partir de deux sources générales d'information.

6.a. La première source est l'information externe. Il s'agit tout d'abord de l'information linguistique, celle qui est portée par l'énoncé. Mais il faut aussi y ajouter l'information qui vient de la situation dans laquelle le compreneur effectue cette compréhension (on parle de "information pragmatique").

6.b. La seconde source est l'information interne. Il s'agit de l'information qui se trouve dans l'esprit du compreneur, cachée aux yeux de tous et même partiellement du compreneur lui-même, sous forme de connaissances de diverses sortes (lexicales, grammaticales, générales); cette information interne est exploitée pour "donner du sens" à l'énoncé. Cette "information interne" est conservée dans la mémoire à long terme où se trouvent emmagasinées les significations de mots, et tout ce qui s'y rapporte. Il faut y adjoindre les connaissances grammaticales du compreneur: sa "compétence linguistique" (= tout ce qu'il a appris de sa langue, connaissances morphologiques, orthographiques, syntaxiques, structurales, connaissances sur les particularités et finesses de sa langue, etc.).

(7) Le compreneur a souvent une connaissance très imparfaite de ses concepts; mais la connaissance qu'il peut avoir de l'usage qu'il fait de ces concepts dans la compréhension est encore plus faible. A première vue, la

compréhension semble bien être une activité consciente. Toutefois, en y regardant de plus près, on voit bien que seul le sens (= la représentation sémantique construite par le traitement de l'énoncé), c'est-à-dire le résultat du processus de compréhension est conscient pour le compreneur. Des prises de conscience partielles du processus ne se produisent pour le compreneur que dans de rares occasions, par ex. lorsqu'il est confronté à des difficultés. C'est sans doute à cause de ce caractère très peu conscient du processus que le sens semble intuitivement "être dans" les énoncés, et qu'il semble pouvoir être immédiatement "saisi".

En réalité, la recherche cognitive a mis en évidence quatre propriétés de la compréhension comme "processus de construction du sens" (voir Le Ny, 116):

- ce processus est automatique
- ce processus est rapide
- ce processus est non conscient
- ce processus est irrépressible (ne peut pas être arrêté)

Conclusions:

Les informations issues des deux grandes sources d'information externe et interne (en bref: l'énoncé et l'esprit du compreneur) forment un ensemble de matériaux assez hétéroclites pour la "construction du sens". Mais il est nécessaire de bien en voir la diversité pour saisir qu'il s'agit véritablement d'une "construction". Ce qu'on peut appeler la "machine à comprendre" ne peut fonctionner qu'en mélangeant, de façon incrémentale, des ingrédients de toute sorte et de toute ampleur, qui devront être extraits (a) de l'énoncé, (b) de la situation d'échange, et (c) de la mémoire du compreneur.

Cette façon d'analyser la compréhension repose donc sur la prise en considération de trois sortes de composants:

- (a) la représentation sémantique terminale en mémoire de travail, qui est l'équivalent psychologique du sens une fois qu'il a été construit;
- (b) les processus qui concourent à cette construction du sens;
- (c) les unités qui servent à cette construction, et qui en constituent les briques.

La contribution propre des mots à la construction du sens se fait par leur signification, c'est-à-dire par l'information sémantique qu'ils véhiculent (nous lui consacrerons un examen détaillé). A propos de ces mots (et de l'ensemble des mots qui constituent le "lexique mental"), on se demandera particulièrement:

- (a) Comment sont les mots dans leur réalité interne? Comment sont-ils "constitués"? En quoi consistent-ils?

- (b) Quelles relations entretiennent-ils les uns avec les autres à l'intérieur du lexique mental? Le lexique mental n'est pas seulement un "réservoir" de mots; c'est surtout un "réseau de représentations" entre lesquelles les relations sont essentielles.
- (c) Comment les esprits utilisent-ils les mots et leurs significations pour faire de la parole et la comprendre?

1. La perception de la parole

1.1. Problèmes généraux

La parole se présente comme un flux sonore continu. La première tâche de l'auditeur (pas facile!) est (a) de segmenter ce flux en unités discrètes et (b) d'identifier ces unités (mots, syllabes, phonèmes).

Si on commence par le niveau le plus élémentaire, on pourrait tenter de déterminer quels sont les indices acoustiques sur lesquels se fonde l'auditeur pour identifier les phonèmes. La phonétique nous donne une masse de données qui permettent de décrire de manière très détaillée les phonèmes (en tous cas, comme unité isolée); toutefois, ces données ne suffisent pas à rendre compte de la perception de la parole continue. Cette dernière, en effet, présente un certain nombre de caractères qui suggèrent l'intervention de processus plus complexes. Voyons quelques-uns de ces caractères.

(a) La rapidité de la parole continue.

La parole continue se déroule couramment à une vitesse de 200 mots par minute; si l'on compte une moyenne de 5 phonèmes par mots, l'auditeur aurait à traiter 1000 phonèmes par minute, c'est-à-dire plus de 15 phonèmes à la seconde.

(b) La variabilité des données acoustiques

La même phrase, qu'on identifie sans peine, correspond à des réalisations sonores très différentes suivant l'âge, le sexe, l'origine du locuteur, selon qu'il parle à voix haute ou basse, la bouche pleine, sous le coup d'une émotion ou non, etc.

(c) Le caractère lacunaire de la parole

Dans la parole continue, une proportion importante de mots (et de phonèmes) est prononcée imparfaitement, voire même pas du tout, sans que cela gêne la compréhension.

Recherche de Pollack & Pickett (1969). Ils firent entendre à des sujets des mots isolés extraits de l'enregistrement d'une conversation spontanée; 47% seulement des mots purent être correctement identifiés.

Conclusion: la perception de la parole continue se fonde bien évidemment sur les données acoustiques, mais elle ne se limite pas à cela. A l'analyse du signal (analyse "de bas en haut", ou bottom-up) s'ajoutent des traitements de "haut en bas" (top-down): l'auditeur anticipe sur ce qu'il entend, et vraisemblablement n'effectue qu'une analyse partielle des données acoustiques (cette analyse ayant pour but de confirmer ou de rectifier ces anticipations). Dans ces conditions, on peut très certainement se demander si

l'unité perceptive de base n'est pas quelque chose de plus grand que le phonème (la syllabe? le mot?); nous reviendrons à cette question en examinant ce que l'on sait aujourd'hui de l'accès lexical.

1.2 La perception des phonèmes

1.2.1 Du son au phonème

Lorsque nous entendons prononcer un mot, les sons perçus constituent une succession de phonèmes distincts (ou plus exactement: une succession de réalisations vocales de phonèmes de la langue). Il semble donc tout à fait raisonnable de croire qu'en découpant convenablement un enregistrement d'un mot, on arriverait à identifier les segments correspondant aux différents phonèmes qui composent ce mot; en découpant par exemple le mot "douter", on obtiendrait quatre sons distincts correspondant aux phonèmes /d/, /ou/, /t/ et /é/.

En fait, il n'en est rien. Voyons quelques observations.

(a) Sur un enregistrement, on peut isoler un fragment de bande magnétique correspondant à la syllabe /dou/; si on raccourcit progressivement l'enregistrement (en prenant un fragment de plus en plus petit), on continuera à entendre /dou/; puis, au moment où la voyelle /ou/ aura disparu, il restera un son qui ne ressemble plus du tout à un son du langage; un sifflement rapide, descendant vers le grave. En réalité, il est impossible d'isoler un son correspondant à /d/.

(b) Si on effectue la même opération avec la syllabe /di/. Cette fois-ci, lorsque la voyelle /i/ a été coupée, ce qui reste est un sifflement, mais montant cette fois-ci vers l'aigu. Conclusion: ce qui est perçu comme un seul et même phonème (une consonne dentale, voisée, non nasalisée /d/) correspond en fait à des signaux acoustiques très différents.

(c) Inversément, un même signal acoustique pourra donner lieu à la perception de phonèmes différents. Par exemple, l'insertion d'un son très bref devant la voyelle /a/ fera que l'auditeur déclare avoir entendu la syllabe /pa/; mais le même son, devant la voyelle /ou/ fera que l'auditeur déclare avoir entendu la syllabe /kou/.

Conclusion: En résumé, il n'y a pas de correspondance terme à terme entre d'une part les données acoustiques objectives, et d'autre part les unités phonémiques que nous percevons.

Du son au phonème

Quelques éléments de phonétique

Le support physique de la parole est une vibration sonore complexe, c'est-à-dire une composition de plusieurs ondes sonores présentant chacune une certaine fréquence et une certaine intensité. On peut représenter la parole sous la forme d'un spectrogramme (voir Figure x ci-dessous):

- + Le temps est représenté en abscisse (de gauche à droite), correspondant au déroulement de la chaîne parolière.
- + La fréquence des vibrations est représentée en ordonnée (du bas = fréquences graves, en haut = fréquences aiguës)
- + L'intensité de chaque fréquence est rendue par le degré de noircissement du spectrogramme (on peut ainsi avoir une image de la répartition de l'intensité sonore sur les différentes fréquences).

Insérer ici figure x sonagramme

On voit sur la figure x que l'intensité se concentre sur certaines bandes de fréquences. Ces concentrations (on les appelle des "formants") définissent les voyelles. Par exemple, en français, la voyelle /i/ se caractérise par trois zones de fréquences renforcées (par trois formants): une première zone aux alentours de 250Hz, une deuxième aux alentours de 2500 Hz et une troisième aux alentours de 3000 Hz.

Les consonnes, elles, sont définies, par les transitions, c'est-à-dire par des changements rapides de fréquence des formants de la voyelle. La transition du second formant, en particulier, selon sa direction (voir figure z) permet de distinguer différentes consonnes (p, b, m, t, d, n, k, g).

On peut ainsi, pour une étude expérimentale, fabriquer des stimuli artificiels dont on pourra faire varier les caractéristiques. (voir ci-dessous, travaux de Liberman et al., perception catégorielle).

1.2.2 Caractères distinctifs et théorie des détecteurs

L'analyse acoustique, nous venons de le voir, ne permet pas de considérer qu'un phonème particulier (en tous cas une consonne) correspond à un objet physique déterminé. L'analyse acoustique met en évidence un ensemble de caractères (par ex. moment d'apparition des formants -apparaissent-ils en même temps, ou avec un certain décalage-, locus de la transition du second formant, etc.) qui, pris tous ensemble, différencient un phonème d'un autre. L'identification d'un phonème pourrait être conçue comme le repérage de la présence ou de l'absence de tel ou tel de ces caractères. Tout se passerait comme si l'identification des phonèmes était le fait d'un ensemble de "détecteurs" spécialisés, fonctionnant en parallèle, et chargés du repérage d'un caractère particulier.

Un ensemble de données, allant dans le même sens, concerne le caractère catégoriel de la perception des phonèmes. Dans une expérience célèbre faite à ce propos (Lieberman et al., 1957), on présentait aux participants des stimuli artificiels, correspondant à des syllabes simples (consonne + voyelle) dont on faisait varier graduellement la transition du second formant (voir figure w1 ci-dessous); ces stimuli étaient identifiés par les participants comme "ba", "da" ou "ga".

Insérer ici figure w1 et w2 Liberman et al., 1957

Deux résultats importants apparaissent:

(a) Dans une première phase de la recherche, on fait entendre un stimulus et le sujet doit dire ce qu'il a entendu (= tâche d'identification). Tous les stimuli sont identifiés sans équivoque par tous les sujets : le passage d'un phonème à un autre se fait de façon abrupte (voir figure w2). Autrement dit, une variation continue du stimulus donne lieu à une perception discontinue, catégorielle.

NB: on a observé un résultat semblable dans de nombreux autres cas. Par exemple, si on accroît graduellement le délai d'apparition du premier formant (voice onset time), on observe une transition brusque de /ba/ à /pa/.

(b) Dans une seconde phase, on présente deux stimuli A et B, puis un troisième C, et on demande aux sujets de dire si le stimulus C est semblable au stimulus A ou au stimulus B (= tâche de discrimination). La discrimination entre deux stimuli voisins (A et B) se fait beaucoup mieux lorsqu'ils se situent près d'une "frontière de phonème" que lorsque tous deux sont perçus comme des

réalisations du même phonème (et cela, même si la différence physique entre les deux stimuli est la même)

Enfin, le même phénomène de perception catégorielle a pu être observé chez des nourrissons de quelques semaines (voir exposés de mars 2005 sur acquisition). Cela semble inviter à considérer les possibles "détecteurs" phonémiques comme relevant d'un équipement inné, spécifique à la perception humaine du langage.

Les choses ne sont cependant pas si simples. Pour différentes raisons: (a) La perception catégorielle ne concerne que les consonnes; rien d'analogue n'apparaît en ce qui concerne les voyelles; (b) des phénomènes de perception catégorielle ont pu être mis en évidence chez des animaux (ex. le chinchilla); (c) l'hypothèse de détecteurs spécialisés ne peut à elle seule rendre compte de la perception de la parole (voir ci-dessous).

1.2.3 Intégration perceptive et contexte.

En fait, l'identification d'un trait phonétique ne se fait pas à partir d'un seul indice acoustique, dont la présence ou l'absence serait une condition nécessaire à l'identification de tel ou tel phonème. C'est une *multiplicité d'indices* qui sont pris en compte.

En outre, l'interprétation de ces indices est dépendante du contexte. En voici quelques exemples:

- Le même son peut être perçu comme deux consonnes différentes suivant la voyelle qui suit. Par exemple, un même son sera perçu comme /p/ ou /k/ selon qu'il est suivi d'un /i/ ou d'un /a/

Son x + i = /pi/

Son x + a = /ka/

- Le même son peut être perçu comme deux consonnes différentes suivant la syllabe qui précède ce son:

/a/ + son x + /a/ = al g a

/ar/ + son x + /a/ = ar d a

Par ailleurs, les indices ne sont pas toujours seulement acoustiques. Une recherche de McDonald & Mc Gurk (1978) montre que des informations visuelles jouent également un rôle. Ils font entendre aux participants de cette recherche différentes syllabes (/ba/, /da/, /ga/, etc.); en même temps, ils présentent sur un écran le visage d'une personne prononçant une syllabe différente de celle qu'ils font entendre. Le résultat, inattendu, est que la plupart du temps, ce qui est perçu, ce n'est ni la syllabe présentée auditivement, ni la syllabe qui correspond à l'information visuelle, mais le

produit d'une intégration des deux sources d'information. Par exemple, si l'information auditive est /ga/, et l'information visuelle /ba/, les sujets entendent /da/.

Conclusion:

Ces différentes observations mettent bien en évidence que la perception des sons de la parole s'appuie sur un ensemble de données très diverses. Ce que les données physiques fournissent, ce ne sont pas des phonèmes, mais une collection d'indices à partir desquels s'effectue un travail d'interprétation dont seul le produit final est accessible à la conscience. Nous verrons brièvement ci-dessous comment on peut concevoir ce "travail d'interprétation".

1.2.4 Interprétation des indices

L'impossibilité de faire correspondre terme à terme les segments de la chaîne sonore à des phonèmes a posé très tôt un sérieux problème aux chercheurs travaillant à la reconnaissance automatique de la parole.

Dès 1962, Halle & Stevens ont proposé une solution possible, qu'ils appellent "analyse par synthèse". L'identification de la parole se ferait à partir d'un processus actif de production. L'auditeur, partant d'une analyse provisoire et partielle du signal, acoustique, produirait "intérieurement" une séquence de phonèmes, qu'il comparerait au message reçu; en d'autres termes, l'auditeur émettrait des hypothèses à partir de ses propres règles de production, et les confronterait aux données acoustiques, jusqu'à ce qu'il aboutisse à un accord satisfaisant.

Une formulation plus précise de ce mécanisme de "synthèse" a été proposée par Liberman et al. (1967) (voir aussi Liberman & Mattingly, 1985) sous la forme d'une "théorie motrice de la perception de la parole". Ce que l'auditeur évoque lorsqu'il écoute un énoncé, ce sont les représentations des mouvements articulatoires qu'il aurait à produire pour émettre ce même énoncé. Ce serait le même mécanisme qui entrerait en jeu dans la production de la parole (qui part de la représentation phonologique de l'énoncé et la traduit en commandes motrices qui produisent des sons) et dans la perception (qui suivrait le chemin inverse). Pour des phénomènes sonores différents (par ex. les diverses réalisations du phonème /d/ selon la voyelle qui suit), ce que l'auditeur perçoit c'est un invariant articulatoire (brève fermeture du canal sonore par la pointe de la langue à la base des dents supérieures, accompagnée d'une vibration des cordes vocales). La perception catégorielle pourrait également se comprendre

de cette façon: au continuum sonore /b, d, g/ par exemple correspondent trois gestes articulatoires distincts.

Ces invariants articulatoires ne correspondent évidemment pas à des mouvements effectifs, mais à de structures générales commandant les mouvements. Autrement dit, percevoir la parole, ce n'est pas la reproduire subvocalement, mais se représenter comment on la produirait si on devait la produire.

L'intérêt de ces modèles est de mettre l'accent sur la "solidarité" entre les processus de perception et les processus de production de la parole; c'est une direction de recherche prometteuse. Par ailleurs, les règles de production que l'auditeur utilise pour construire les "hypothèses" qu'il confronte aux données acoustiques de l'input ne sont pas forcément limitées aux règles articulatoires: elles peuvent mettre en jeu des niveaux plus élevés; lexical, syntaxique, sémantique. C'est ce que nous allons voir ci-dessous.

1.3 La perception de la parole continue

1.3.1 Le rôle du contexte

Une recherche classique de Miller & Isard (1963) montre bien le rôle joué par l'organisation linguistique dans la perception de la parole. Les participants à cette recherche entendent des séquences de mots présentés dans un bruit masquant, et doivent simplement répéter ce qu'ils ont entendu. Il y avait 4 types de séquences:

- (1) Des phrases grammaticalement correctes et sémantiquement acceptables; par ex. "Accidents kill motorists on the highways" (Des accidents tuent des automobilistes sur les autoroutes)
- (2) Des phrases grammaticalement correctes, et sémantiquement inacceptables; par ex. "Accidents carry honey between the house" (Des accidents apportent du miel entre la maison)
- (3) Des séquences obtenues en brouillant l'ordre des mots des phrases de type (1); par ex. "Accidents motorists on kill the highways".
- (4) Des séquences obtenues en brouillant l'ordre des mots des phrases de type (2); par ex. "Accidents honey between carry the house".

Les résultats sont très clairs. Il n'y a pratiquement pas d'erreur dans la répétition des phrases (1); par contre, il y a un maximum d'erreurs pour les phrases (4); les séquences de type (2) et (3) donnent lieu à des performances intermédiaires. La structure syntaxique et la cohérence sémantique jouent un rôle facilitateur, et permettent à l'auditeur de reconstituer les informations phonétiques masquées par le bruit.

L'auditeur n'a pas conscience de cette "reconstitution". Le même stimulus sera d'ailleurs perçu différemment selon le contexte. Ainsi, dans la phrase "It was found that the *eel was on the axle" (On s'aperçut que la *eel était sur l'essieu), les sujets entendent distinctement "wheel" (roue). Mais si on leur présente "It was found that the *eel was on the orange", ils entendent "peel" (pelure); et si on leur présente "It was found that the *eel was on the table", ils entendent "meal" (repas).

It was found that the *eel was on the axle" (wheel).

It was found that the *eel was on the orange" (peel).

It was found that the *eel was on the table" (meal).

Autres observations. Cole et al. (1978) font entendre des histoires où certains phonèmes ont été remplacés par d'autres; les participants à la recherche doivent détecter ces erreurs le plus vite possible. Les résultats (pourcentage d'erreurs détectées, rapidité de la détection) diffèrent très nettement selon que le phonème remplacé est situé au début ou à la fin d'un mot, sur une syllabe accentuée ou non, dans un mot plus ou moins prévisible d'après le contexte, etc. D'une façon générale, plus le phonème déformé est prévisible, plus la déformation tend à passer inaperçue; mais, si elle est détectée, elle le sera alors plus rapidement.

Toutes ces observations mettent en évidence une interaction entre deux types de traitement: un traitement basé sur l'analyse des données acoustiques (traitement de type "bottom-up", de bas en haut), et un traitement basé sur des anticipations liées au contexte (traitement de type "top-down", de haut en bas). Ces observations conduisent aussi à poser la question suivante: l'identification des phonèmes est-elle bien le phénomène premier, simple et élémentaire, sur lequel se fondent les opérations de niveau supérieur? Le phonème est une unité linguistique incontestable; mais a-t-il une "réalité cognitive"? (c'est-à-dire: correspond-il à quelque chose qu'utilise notre cerveau lorsqu'il traite du langage?).

1.3.2 Quelle est l'unité perceptive de base?

Nous sommes bien évidemment capables de reconnaître un phonème, et par exemple capables de détecter l'élément commun aux mots "bas" et "bout". Mais s'agit-il d'un donné perceptif immédiat? On sait que la conscience phonologique chez l'enfant, établie par son aptitude à segmenter en phonèmes, n'apparaît que vers 6 ans, c'est-à-dire à l'âge où se fait l'apprentissage du langage écrit. L'identification des phonèmes pourrait n'être que le résultat

d'un apprentissage et d'une analyse de données perceptives d'un niveau supérieur.

A l'appui de cette idée, on peut citer les résultats obtenus dans des tâches de détection de phonèmes ("phoneme monitoring").

Le sujet, tout en écoutant une phrase, doit appuyer sur une touche lorsqu'il entend, par exemple, un mot commençant par /b/. Le système étant à capacité limitée, on peut s'attendre à ce que l'identification du phonème soit retardée (donc, temps de réponse plus long) si la partie de l'énoncé qui précède immédiatement le phonème demande un traitement plus complexe.

En utilisant ce type de tâche, on a présenté une liste de syllabes, en demandant aux participants de détecter soit une syllabe particulière (par ex. pousser sur un bouton dès que l'on entend /ba/), soit seulement le phonème initial de cette syllabe (par ex. /b/). Le résultat obtenu est très surprenant: les syllabes sont détectées plus vite que les phonèmes. Les chercheurs en ont conclu que le phonème n'a pas de réalité perceptive, et que l'unité de perception est la syllabe, le phonème n'étant identifié que par analyse de la syllabe (on identifierait d'abord la syllabe /ba/ et c'est en analysant cette syllabe qu'on isolerait le /b/).

D'autres observations invitent même à considérer le mot lui-même comme unité perceptive de base. Par exemple, on présente aux participants une liste de mots bi-syllabiques (ex. bateau, palais, vélo, mouton, poisson, page, dormir, etc), et on leur demande de détecter le plus rapidement possible soit un mot, soit une syllabe, soit un phonème. Le temps de détection est plus court pour la première syllabe que pour le phonème initial (= on détecte plus vite /pa/ que /p/), mais il est encore plus court pour le mot lui-même (= on détecte plus vite /palais/ que /pa/). Et cet effet est d'autant plus marqué que le mot est davantage prévisible à partir du contexte.

- Les visiteurs admiraient les photos du palais
- Le roi a reçu ses invités dans les jardins du palais

Toutefois, on a également montré que la détection de phonèmes peut très bien s'effectuer sur des pseudo-mots (qui, par définition, ne peuvent se trouver en mémoire).

1.3.3 Conclusions

Tout ce qui précède suggère l'idée d'un traitement parallèle, à plusieurs niveaux, avec interaction possible entre niveaux.; l'un ou l'autre de ces niveaux fournissant (selon la tâche, les stratégies du sujet, la structure de la langue, etc.) l'information nécessaire à la détection.

Voir figure 4-8, Carroll, p.98

Peut-on envisager que la perception de la parole constitue un mode de perception spécifique, propre à l'espèce humaine, faisant appel à des mécanismes propres? Il est possible que le cerveau humain comporte des dispositifs spécialisés d'analyse du signal acoustique, propres à la fonction du langage, et faisant partie du bagage héréditaire de l'espèce humaine. Cette hypothèse, toutefois, est encore très problématique.

L'originalité de la perception de la parole tient plus à sa complexité qu'à sa nature même. Ce qui la caractérise, c'est d'une part la multiplicité des niveaux de traitement, et d'autre part l'importance des traitements de type Top-Down.

2. L'accès lexical

2.1 Le lexique interne

L'idée de base sur laquelle repose toute étude du lexique mental est que tous les êtres humains ont dans leur esprit/cerveau (dans leur mémoire à long terme), à partir de leur deuxième année, un très grand nombre de "mots", ou plus exactement un très grand nombre de "représentations" de mots. Pour reconnaître ou produire un mot, il faut en effet que ce mot ait été enregistré en mémoire. Ces représentations de mots constituent une part si importante de la mémoire à long terme des individus qu'on lui attribue une dénomination spécifique: le "lexique mental" ou "lexique interne".

Les recherches sur le lexique mental s'appuient sur l'hypothèse qu'il possède des propriétés générales, communes à tous les locuteurs. Bien sûr, tous les lexiques mentaux diffèrent les uns des autres (selon la ou les langues parlées, selon le milieu social, selon l'âge, selon les caractéristiques personnelles). Mais on fait l'hypothèse qu'en dépit de ces différences individuelles, tous les lexiques mentaux présentent des propriétés générales communes. Il existe un socle commun à tous les locuteurs humains, sous forme de structures fondamentales du langage, et de modes de fonctionnement de ces structures.

Parler de "lexique mental" implique que l'on accepte, à côté de la notion de "mot" comme réalité linguistique, la notion de "mot mental", c'est-à-dire la notion de "représentation de mot". Le locuteur ordinaire, et même le linguiste qui se propose d'étudier objectivement les faits de parole, se trouvent en présence de "mots linguistiques", oraux ou écrits, qui sont des réalités physiques observables. On peut les décrire et les caractériser de diverses façons. Le psychologue par contre s'intéresse aux "mots dans la tête", c'est-à-dire à une certaine sorte d'unités cognitives, dont on doit postuler l'existence pour rendre compte du fonctionnement du langage. Ces "mots mentaux" correspondent, dans l'esprit des locuteurs, aux "mots physiques" qu'ils entendent ou prononcent, qu'ils lisent ou écrivent.

On peut envisager ce lexique sur le modèle du dictionnaire, comme constituant un "répertoire de représentations de mots", conjointement de leurs formes et de leurs significations. Le lexique se présenterait dans ce cas comme une liste ordonnée d'"entrées lexicales"; chaque entrée comporterait l'ensemble des informations nécessaires pour identifier, comprendre et utiliser l'unité considérée.

Trois types d'informations, au moins, devraient caractériser chaque entrée lexicale:

- + (1) la forme phonologique (acoustique, articulatoire, graphique) du mot.
- + (2) ses propriétés syntaxiques: catégorie (nom, verbe, adjectif, etc.), genre (masculin, féminin), nombre (singulier, pluriel), etc. définissant ses possibilités d'utilisation dans une phrase.
- + (3) sa signification, et sans doute ses conditions d'utilisation selon la situation (mot familier, vulgaire, etc.).

Les questions principales que nous devons nous poser pour comprendre le lexique mental portent d'une part sur le contenu et l'organisation de ce lexique (ci-dessous questions 1 et 2), et d'autre part sur les processus d'accès et d'utilisation de ce lexique (ci-dessous question 3):

- + (1) Comment sont les mots dans leur réalité interne? Comment sont-ils constitués?
- + (2) Quelles relations entretiennent-ils les uns avec les autres à l'intérieur du lexique mental? Quelle est la structure sémantique de ce lexique mental?
- + (3) Comment les esprits utilisent-ils les mots et leurs significations pour faire de la parole et la comprendre? Que sait-on des processus d'accès au lexique mental: s'agit-il de processus passifs et automatiques, ou actifs? De processus autonomes (se déroulant indépendamment des autres traitements éventuellement en cours) ou interactifs (se déroulant en interdépendance avec des processus qui s'effectuent simultanément à d'autres niveaux)?

2.2 La tâche de décision lexicale

Pour aborder ces problèmes, l'introspection n'est d'aucun secours. L'accès lexical est en effet un processus très rapide, qui se déroule de manière inconsciente. Pour l'étudier, il faut donc avoir recours à des techniques expérimentales permettant de savoir comment se déroule l'accès à la représentation de la forme des mots, et à travers elle, à la signification. De toutes les mesures "on line" (permettant de saisir les processus au cours même de leur déroulement, en temps réel), la plus utilisée est celle qu'offre la tâche de "décision lexicale".

La procédure expérimentale consiste à présenter des stimuli, et à demander aux participants de décider pour chaque stimulus si il est ou non un mot. Dans la version écrite, plus facile à mettre en oeuvre que la version orale, on présente successivement des séquences de lettres (des "chaînes de caractères") sur un écran, et on demande aux participants d'appuyer le plus vite possible sur un bouton "oui" s'il s'agit d'un mot (par exemple "chapeau"), et sur un bouton "non" si ce n'est pas le cas (par exemple, "chadeau", "pacheau",

ou "cmvhbn"). Pour des séquences du type "cmvhbn", on parle de "non-mots"; pour des séquences de type "chadeau" (qui diffère d'un mot réel par une ou deux lettres), on parle de "pseudo-mots". La réponse est généralement donnée par un appui sur une touche d'un clavier d'ordinateur; une des touches du clavier est baptisée "oui", et une autre touche "non".

Cette procédure est à peu près exclusivement utilisée pour mesurer la latence de la réponse, c'est-à-dire le temps qui s'écoule entre l'apparition du stimulus (apparition de la séquence de lettres sur écran) et l'instant où la touche du clavier est enfoncée par la réponse du participant. Cette durée est le "temps de décision lexicale": elle est de l'ordre de quelques centaines de millisecondes (entre 500 et 700 millisecondes), et varie en fonction de divers facteurs (voir ci-dessous).

2.3 Deux phénomènes de base: fréquence et amorçage.

Dès les premiers travaux sur l'accès lexical, deux phénomènes très généraux ont pu être mis en évidence: (1) l'effet de la fréquence d'usage des mots, et (2) l'effet d'amorçage.

2.3.1 Effet de fréquence d'usage.

On conviendra sans peine que les mots de la langue (français, anglais, vietnamien, etc.) sont employés de façon plus ou moins fréquente. Cette fréquence d'usage ne peut évidemment être définie qu'approximativement; elle dépend notamment du milieu social, culturel, professionnel, du type de conversation ou de lecture, etc. On peut estimer la fréquence d'usage en interrogeant diverses bases de données (surtout constituées de corpus de langage écrit, oeuvres littéraires, articles de journaux, etc.), ce qui permet de l'exprimer en nombre d'occurrences du mot parmi 100 millions.

Si on présente à un groupe de participants une série de mots qui ont des fréquences variées dans la langue, et qu'on leur demande d'estimer le sentiment de familiarité que chacun de ces mots a pour eux, par exemple en utilisant une échelle allant de 1 ("ce mot m'est très peu familier") à 7 ("ce mot m'est très familier"), on fait deux constatations: (a) tout d'abord, cette tâche est facile et donne lieu à un très grand accord entre les participants; et surtout (b) il y a une très forte corrélation positive entre le sentiment de familiarité et la fréquence des mots dans la langue. Autrement dit: les participants n'ont aucune connaissance des fréquences objectives, mais leur sentiment de familiarité en est un excellent reflet.

De nombreuses expériences ont montré que le temps nécessaire pour traiter mentalement un mot, le percevoir, effectuer une décision lexicale à son sujet, le prononcer, dénommer un objet, dépend de la fréquence (ou familiarité) de ce mot. Dans la décision lexicale par exemple, si on présente,

parmi des pseudo-mots, des mots de diverses fréquences d'usage, bien contrôlés par ailleurs sur de nombreuses autres caractéristiques (longueur, composition syllabique, etc.), on constate que les temps de réponse "oui" varient en fonction de la fréquence ou familiarité du mot. Plus un mot est fréquent, plus son accès dans le lexique interne est rapide.

2.3.2 Effet d'amorçage

Le second phénomène important est celui de l'amorçage: l'accès à un mot Z est facilité (et donc est plus rapide) si ce mot Z a été précédé par un autre mot M dont la forme ou la signification sont proches de la forme ou de la signification du mot Z.

Exemples :

DOCTEUR	DOCTEUR
MALADE	DOCTEUR
PAIN	DOCTEUR

MALADE	BEURRE
PAIN	BEURRE

Le lexique n'est pas une accumulation de mots; c'est un système d'unités reliées entre elles. Cet aspect des choses apparaîtra lorsque nous parlerons des significations de mots, et des liaisons qui existent entre ces significations de mots); mais les formes des mots sont également liées entre elles par de multiples liaisons, et on peut dire qu'elles forment aussi un réseau, un réseau de formes.

2.4 Que contient le lexique?

A première vue, la question est simple, et la réponse semble évidente: le lexique contient les mots de la langue. Il suffit toutefois de considérer quelques questions pour se rendre compte que les choses ne sont pas aussi simples:

- Qu'est-ce qu'un mot?
- Que se passe-t-il pour les mots polysémiques?
- N'y-a-t-il pas plusieurs lexiques?

(a) Qu'est-ce qu'un mot?

D'un point de vue linguistique, on sait que la plus petite unité dotée de sens n'est pas le mot mais le morphème. Le lexique interne contient-il tous les mots de la langue (par ex. aimable, déplaçable, in mangeable, introuvable, insupportable, lavable, mangeable, etc.) ou seulement les morphèmes qui les

constituent (par ex. aim-, déplac-, in-, mang-, trouv-, support-, lav-, mang-, -able, etc.), et qui peuvent entrer dans toutes sortes de combinaisons.

Très concrètement: toutes les formes d'un verbe sont-elles stockées en mémoire (parle, parlais, parlerai, parlé, parlions, parliez, parlera, parlerons, parlerez, etc.) ou seulement le radical de ce verbe ("parl-"), et les désinences (-ais, -erai, -é, -ions, -iez, -era, etc.) qui peuvent être utilisées avec un grand nombre d'autres verbes?

(b) Et les mots polysémiques?

Hypothèse 1: un mot comme "avocat" comporterait deux entrées lexicales (avocat1= juriste, et avocat2= fruit) dont l'accès serait commandé par le contexte? Selon le contexte, ce serait l'entrée 1 ou l'entrée 2 qui serait activée.

Hypothèse 2: nous n'aurions en mémoire qu'un seul mot, dont la signification (avocat1 ou avocat2) serait sélectionnée après l'accès lexical?

Les données expérimentales sont plutôt en faveur de l'hypothèse 2. Notamment l'expérience de Swinney (1979) (voir détails de cette recherche ci-après).

Expérience de Swinney (1979)

Les participants écoutent des petits textes; tout en écoutant ces textes, ils regardent un écran sur lequel apparaît de temps en temps une séquence de lettres; lorsqu'une séquence de lettres apparaît, les participants doivent dire le plus vite possible si c'est un mot ou non.

Dans chaque texte, on insère une phrase expérimentale qui contient un mot polysémique dont un sens particulier est favorisé par le contexte qui précède. Par exemple: un texte comprenait la phrase "Le boulanger emballa la baguette qu'il venait de sortir ..." ("baguette1= pain allongé", "baguette2= petit bout de bois"; le contexte favorise le sens baguette1). On s'arrange pour faire apparaître sur l'écran une séquence de lettres:

(a) soit immédiatement après l'audition du mot polysémique baguette: quand le sujet entend "-ette" (" ... emballa la baguette...), il voit apparaître sur écran une séquence de lettres, et doit dire si cette séquence constitue un mot ou non;

(b) soit 3 syllabes après la fin du mot polysémique baguette: quand le sujet entend "-ait" (" emballa la baguette qu'il venait ...), il voit apparaître sur écran une séquence de lettres, et doit dire si cette séquence constitue un mot ou non;

La séquence de lettres est:

- (a) un mot associé au sens favorisé par le contexte: **Blé**
- (b) un mot associé au sens non favorisé par le contexte: **Bois**
- (c) un mot neutre: **Gare**

Le temps de décision lexicale pour le mot neutre (Gare) ne devrait pas être affecté par le contexte; ce temps servira de repère. On dira qu'un des sens de baguette (baguette1 ou baguette2) est activé par le contexte si le temps de décision lexicale pour le mot qui lui est associé est plus bref que pour gare.

	Temps de décision lexicale		
	Blé	Bois	Gare
Immédiatement après le mot ambigu:	<u>890</u>	<u>900</u>	960
Trois syllabes après le mot ambigu:	790	<u>860</u>	<u>860</u>

Immédiatement après avoir entendu "baguette", la décision pour "blé" et pour "bois" est plus rapide que la décision pour "gare"; mais il n'y a pas de différence entre le temps de décision pour "blé" et le temps de décision pour "bois".

Après trois syllabes, seule la décision pour le mot (blé) associé au sens favorisé par le contexte est plus rapide que la décision pour le mot neutre (gare).

Conclusion: initialement, les différents sens d'un mot ambigu sont activés; ultérieurement, le contexte sélectionne le sens adéquat. Il y aurait donc d'abord accès lexical à une entrée lexicale unique (accès à l'entrée lexicale "baguette"), puis après, sélection d'un sens particulier permettant l'intégration au sens de la phrase.

(c) Un seul ou plusieurs lexiques internes?

On peut se demander si les mots fonctionnels (articles, pronoms, prépositions, conjonctions, etc.) sont enregistrés de la même façon, et dans le même lexique, que les mots pleins (noms, verbes, adjectifs, etc.).

Est-ce le même lexique interne qui fonctionne pour le langage oral et pour le langage écrit, ou a-t-on deux lexiques différents? N'existe-t-il pas un lexique interne spécifique pour l'écrit (un lexique dont les entrées lexicales se présenteraient sous forme graphique)?

Ne doit-on pas considérer que production et compréhension font appel à deux lexiques internes distincts? (Cfr différenciation des aphasies selon qu'elles touchent l'expression (Broca) ou la réception (Wernicke) + travaux de neurolinguistique).

2.5 Le(s) mécanisme(s) de l'accès lexical

Nous devons également nous demander quel est le mécanisme (ou le processus) par lequel s'effectue l'accès lexical. Quel est le mécanisme par lequel, à partir d'un stimulus auditif ou visuel, nous retrouvons en mémoire un mot, avec toutes les informations qui y sont liées? On peut concevoir ce processus de différentes façons.

2.5.1 Mécanisme de recherche séquentielle.

On a tout envisagé l'accès au lexique sur le modèle de la consultation d'un dictionnaire. Pour chercher un mot, nous feuilletons le dictionnaire, nous localisons la page où le mot doit vraisemblablement se trouver, puis nous parcourons la liste des mots de cette page jusqu'à ce que nous ayons trouvé le mot cherché. On procède donc par une recherche séquentielle active.

Exemple: le modèle de Forster.

Dans ce modèle, le lexique interne est constitué de quatre fichiers (voir figure). Un fichier central ("master file") contenant l'ensemble des informations (morphologiques, syntaxiques, sémantiques) sur chaque mot; on a accès à ce fichier central via des fichiers périphériques:

- (a) un premier fichier comporte des entrées correspondant à la forme graphique des mots (et est utilisé dans le traitement du langage écrit);
- (b) un deuxième fichier comporte des entrées correspondant à la forme phonologique des mots (et est utilisé dans le traitement du langage oral);
- (c) un troisième fichier comporte des entrées organisées sur une base syntaxico-sémantique (et est utilisé pour la production de la parole)

Chacun de ces fichiers périphériques est organisé sous forme de "casiers": à l'intérieur de chaque casier, les mots sont classés selon leur fréquence. Sur la base d'informations sensorielles (acoustiques ou visuelles), l'auditeur (ou le lecteur) va donc limiter sa recherche à un "casier", qu'il parcourra dans l'ordre jusqu'à ce qu'il ait trouvé le mot correspondant à ce qu'il a perçu; il sera alors renvoyé au fichier central où il trouvera les informations concernant ce mot, ses associations avec d'autres, etc.

2.5.2 Mécanisme d'activation automatique.

Une autre hypothèse serait qu'il n'y a pas de recherche, mais simplement une activation automatique des mots (ou plus exactement des dispositifs producteurs de mots, qu'on a appelé des "logogènes") par les informations entrant dans le système: dès qu'un certain seuil d'activation est atteint, le "logogène" entre en action et le mot correspondant est produit ou identifié. On peut comparer ce mécanisme à ce qui se passe pour un verre de cristal qui se met à résonner quand un instrument de musique (ou une voix) produit, dans le voisinage, une certaine note avec une certaine intensité.

Exemple: le modèle de Morton.

Dans ce modèle, à chaque mot du lexique interne correspond un "logogène" (dispositif producteur de mot, dispositif qui à un certain degré d'activation produit la représentation consciente d'un mot).

Les logogènes sont activés par les informations fournies au système: informations auditives ou visuelles, mais aussi informations contextuelles, liées au contexte linguistique, à la situation et aux connaissances générales du sujet. Toutes ces informations interviennent à la fois et cumulent leurs effets: le logogène se déclenche lorsqu'il a atteint un certain seuil d'activation, et le mot devient disponible. Après son déclenchement, le logogène ne retrouve que peu à peu son état initial: il reste encore un certain temps partiellement activé. Ce qui permet d'expliquer l'effet de fréquence (un logogène fréquemment utilisé n'est jamais totalement désactivé, et sera donc plus rapidement déclenché), et l'effet d'amorçage (l'activation d'un logogène entraîne une activation partielle des autres logogènes qui lui sont "proches").

Contrairement au modèle de recherche séquentielle (à la Forster), on voit que le modèle d'activation automatique (à la Morton) suppose une interaction permanente entre toutes les sources d'information et tous les niveaux de traitement. En outre, l'activation des logogènes est un mécanisme automatique, et donc passif.

2.5.3 Caractère interactif des processus en jeu.

D'autres recherches (par ex., Marslen-Wilson, 1985) ont clairement mis en évidence l'influence du contexte sur les processus d'accès lexical, et dès lors ont bien souligné le caractère interactif de ces processus. Voyons rapidement quelques résultats significatifs.

(a) L'identification d'un mot, présenté dans un contexte cohérent, se fait non seulement très vite, mais très tôt par rapport à la durée du mot lui-même: le mot est généralement reconnu avant d'avoir été complètement entendu.

Dans une tâche de "shadowing", les sujets peuvent répéter le texte qu'ils entendent avec un décalage qui n'est que de 250 millisecondes. On observe un temps moyen de réponse de même ordre (275 millisecondes) quand les sujets doivent détecter un mot (qu'on leur a préalablement indiqué) dans une phrase qu'on leur fait entendre. Or, la durée moyenne d'un mot, dans ces travaux, est de 370 à 400 msec. Lorsqu'ils répondent, les sujets ne peuvent donc avoir entendu que les deux premiers phonèmes.

(b) Cette identification du mot dépend du contexte. Quand on fait varier l'information syntaxique et/ou sémantique dont dispose le sujet, le temps d'identification varie en conséquence.

Si, au lieu de présenter le mot dans une phrase normale, on présente aux sujets des séquences de mots en désordre, le temps moyen de réponse passe de 275 msec à 380 msec. L'effet de facilitation augmente à mesure que le mot-cible (à détecter) est placé plus tard dans la phrase; mais cet effet de facilitation apparaît dès le début de la phrase si celle-ci est précédée d'une autre phrase qui lui procure un contexte cohérent.

(c) Enfin, la reconnaissance d'un mot est lié à la composition du lexique lui-même: le mot est identifié dès que le sujet a recueilli assez d'informations pour pouvoir le distinguer des autres mots possibles.

On peut montrer cela par une mesure du temps de décision lexicale pour des "non-mots". Par exemple, soit les deux non-mots suivants en français: "xomène" et "révolque". Dans le premier cas (xomène), dès le second phonème (o), on peut décider qu'il ne s'agit pas d'un mot (aucun mot français ne commence par "xo-"). Dans le second cas (révolque) par contre, il faut attendre le sixième phonème (q) pour décider qu'il ne s'agit pas d'un mot (plusieurs mots français commencent en effet par "révol-")

Pour rendre compte de tous ces résultats, Marslen-Wilson a proposé un modèle interactif du traitement du langage. Dès le début de la perception d'une phrase, tous les niveaux de traitement (phonologique, morphologique, lexical, syntaxique, sémantique) commencent à fonctionner en parallèle; et ces divers niveaux de traitement sont en interaction: chaque niveau est affecté par les informations provenant non seulement du niveau immédiatement

inférieur, mais aussi des niveaux supérieurs. Cela permet un fonctionnement optimal et économique du système: l'interprétation du message se fait dès que les informations disponibles (quelle que soit leur source) permettent de l'effectuer.

En ce qui concerne l'accès lexical, l'identification des mots se fait par élimination progressive. Dès le début d'un mot, une "cohorte" de mots compatibles avec les informations disponibles est évoquée; à mesure que l'information s'enrichit, le nombre des "candidats" possibles se restreint; et quand il ne reste plus qu'un seul mot possible (ce qui peut, on l'a vu, se passer très rapidement), c'est lui qui est retenu.

3. Du mot à la signification

Nous avons jusqu'ici traité du lexique mental en nous intéressant surtout aux représentations de la forme des mots, aux représentations par lesquelles s'effectue l'accès au lexique. Nous devons maintenant nous intéresser à la signification des mots. Règlons tout d'abord un problème de terminologie. Dans ce qui suit, nous considérerons qu'il existe dans l'esprit des locuteurs (c'est-à-dire dans leur mémoire à long terme) des entités cognitives que nous désignerons en parlant de "signification de mot", "représentation sémantique", "concept" ou "catégorie". Ces différentes expressions visent une même réalité cognitive que l'on cherche à caractériser de manière scientifique.

Deux grands types de problèmes doivent être envisagés:

(a) Quelle est la nature de ces significations?

Nous nous demanderons tout d'abord comment la signification d'un mot est représentée mentalement. Ce qui revient à se demander comment les "catégories cognitives" sont faites dans l'esprit des individus, comment elles sont constituées. Cela conduit à plusieurs autres questions que nous ne ferons qu'effleurer:

- + Comment les catégories se forment-elles, puis évoluent-elles au cours de la vie?
- + Existe-t-il des bases innées aux catégories? Si oui, lesquelles?
- + Comment les catégories apparaissent-elles chez le bébé, et comment se développent-elles chez l'enfant et l'adolescent?
- + Comment, plus tard, de nouvelles catégories, de nouveaux concepts, sont-ils appris, notamment par l'intermédiaire du langage?

(b) Comment les significations de mots sont-elles organisées en mémoire?

Poser cette question, c'est se poser la question de l'organisation de la mémoire sémantique. C'est se demander comment les significations de mots fonctionnent au cours des activités psychologiques qui les mettent en oeuvre (production ou compréhension du langage).

3.1 Comment la signification d'un mot est-elle représentée?

Il est facile de dire que l'on connaît ou non la signification d'un mot; il est beaucoup moins facile de préciser ce qu'on entend par là. Comme nous allons le voir, rapidement, il peut s'agir d'un ensemble de traits sémantiques (la sémantique componentielle), une représentation prototypique (théorie du prototype), un réseau de connaissances (toutes les connaissances liées à un

mot particulier), une procédure (le sens d'un mot, c'est son usage, l'utilisation qu'on en fait dans un contexte particulier).

3.1.1 La conception componentielle

On peut envisager de définir la signification d'un mot comme l'ensemble des "traits" par lesquels il se différencie de la signification des autres mots de la langue. Soient par ex. les traits "Vivant/Non vivant", "Humain/Non humain", "Mâle/Non mâle", "Adulte/Non adulte": selon l'assemblage, cela donnerait "Homme", "Femme", "Garçon", "Fille", "Veau", "Vache" etc.

Des observations expérimentales sont compatibles avec cette façon de concevoir la signification d'un mot.

(a) Observations classiques concernant les associations verbales. Par exemple, la fréquence élevée des associations par contraste pourrait s'expliquer par le fait qu'on ajoute ou qu'on supprime un ou plusieurs traits: Homme - Femme, Garçon - Fille, etc.

(b) Observations relatives à l'acquisition du langage. On peut considérer le développement des significations dans le vocabulaire enfantin comme l'acquisition progressive de traits sémantiques (E. Clark). En accord avec cela, on peut citer:

- + Les surextensions chez le jeune enfant: "chien" est utilisé pour désigner tous les animaux à quatre pattes, parce que ce mot n'a pas encore tous les traits sémantiques qui permettront de limiter son usage.
- + On constate chez les jeunes enfants des confusions entre des termes qui ne diffèrent que par un seul trait. Par ex. confusion entre "more" et "less", "avant" et "après", "aller" et "venir", etc.
- + L'ordre d'acquisition: le couple "grand/petit" est acquis plus tôt que "haut/bas" qui comporte le trait supplémentaire [+ vertical].

3.1.2 La conception propositionnelle

D'autres auteurs renoncent à l'idée d'une décomposition en traits élémentaires, et définissent la signification d'un mot comme une liste de propriétés. L'unité sémantique de base n'est donc plus conçue comme un trait sémantique, mais comme une proposition (c'est-à-dire comme une relation entre deux ou plusieurs termes).

Exemples de propositions:

ABOYER (Chien)

CUIRE (Boulangier, Pain)

BRILLER (Etoile, Ciel)

La signification d'un mot peut être représentée sous forme d'une liste de propositions: cette liste représente l'ensemble des connaissances attachées à ce mot.

Selon cette conception, la représentation des significations en mémoire correspond davantage à une encyclopédie qu'à un dictionnaire.

Un intérêt de cette conception est le suivant: le même format propositionnel permet de décrire la signification de mots et la signification de phrases:

Soit le mot "Boulangier":

Sa signification est représentée par <CUIRE (Boulangier, Pain)>

Soit la phrase "Kim a vu la fusée partir"

Sa signification est représentée par deux propositions:

p1 VOIR (Kim, p2)

p2 PARTIR (fusée)

Cette représentation propositionnelle permet de rendre compte de l'organisation des textes; on peut en effet décomposer n'importe quel discours en un ensemble de propositions hiérarchisées les unes par rapport aux autres.

Illustration: cfr figure Texte de Queneau, in JC p. 198.

3.1.3 La conception prototypique

Les concepts auxquels renvoient les mots d'une langue n'ont pas le caractère précis et bien défini des concepts scientifiques: les catégories qu'ils constituent ont des contours souvent flous, et l'appartenance à ces catégories n'est pas du genre "tout ou rien", mais plutôt du genre "plus ou moins".

Par exemple, une table ou un lit sont des meubles (= appartiennent à la catégorie "mobilier"); mais qu'en est-il d'une lampe de chevet? D'un téléviseur? D'une poubelle de bureau?

L'appartenance n'est pas du type tout ou rien. Par exemple, une hirondelle est plus un oiseau qu'un héron; une truite est plus un poisson qu'un

hippocampe; une pomme est plus un fruit qu'une noisette; et, dans une certaine mesure, on peut considérer la baleine comme un poisson; etc.

Ce que ces exemples montrent c'est que, pour une catégorie donnée (mobilier, poisson, fruit, etc.), certains exemplaires sont plus typiques que d'autres. Certains auteurs (E. Rosch, 1981) pensent que la représentation mentale d'un concept serait celle d'un représentant privilégié: le prototype. L'appartenance à une catégorie serait fonction du degré de ressemblance au prototype.

Des données expérimentales nombreuses ont été recueillies à ce sujet. Quand on demande à des sujets de citer par exemple tous les noms d'oiseaux qu'ils connaissent (ou tous les noms de poisson, de vêtement, de moyen de transport, etc.), l'ordre et la fréquence avec lesquels sont évoqués les exemplaires d'une catégorie donnée permettent de dégager des échelles de typicalité. Dans des tâches de vérification de phrases, les temps de réponse sont plus courts si la phrase porte sur un exemplaire typique de la catégorie ("Une pomme est un fruit") que sur un exemplaire atypique ("Une noix est un fruit"). Par ailleurs, les exemplaires typiques sont plus facilement mémorisés, etc.

Cette conception souligne le fait que la signification des mots ne se réduit pas à quelque chose de précis et de bien défini, mais consiste plutôt en un "air de famille" entre des objets qui peuvent par ailleurs être très différents. Il reste toutefois, pour l'instant, que les notions de typicalité et de prototype n'ont qu'une valeur descriptive, et que leur statut cognitif reste à déterminer.

3.2 Comment la mémoire sémantique est-elle organisée?

Quelle que soit la réponse que l'on donne à la question de savoir de quelle nature est la représentation de la signification (traits sémantiques, propositions, prototypes), il faut également se demander comment ces significations sont organisées en mémoire, et par quels processus elles sont évoquées. Autrement dit, nous devons examiner l'organisation et le fonctionnement de la mémoire sémantique.

Un premier groupe de théories s'intéresse principalement aux significations de mots (significations lexicales); d'autres théories, plus ambitieuses, portent sur le problème plus général de l'organisation et du fonctionnement des connaissances.

3.2.1 Les significations lexicales

Comment les significations de mots sont-elles stockées en mémoire? Plusieurs théories ont été avancées à ce sujet:

- (a) Selon certains auteurs (voir ci-dessous Modèle de Collins & Quillian), les significations sont stockées sous forme d'un ensemble de relations que chaque mot entretient avec d'autres mots
- (b) Selon d'autres auteurs (voir ci-dessous Modèle de Rips et al.), les significations sont stockées sous forme de listes de propriétés enregistrées avec le mot.

Modèle en réseau: Collins & Quillian

Le premier modèle de mémoire sémantique a été proposé en 1966 par Collins et Quillian. Selon ce modèle, la mémoire sémantique se présente comme un vaste réseau constitué de noeuds et d'arcs reliant les noeuds entre eux. Les noeuds sont des "concepts". Les "arcs" correspondent aux relations (aux "liaisons") entre concepts; ils sont bidirectionnels (un arc part d'un concept Z vers un autre, et arrive de cet autre vers Z). Au départ d'un concept, on trouve donc un certain nombre d'arcs qui mènent à d'autres concepts, d'où partent de nouveaux arcs menant à d'autres concepts encore, et ainsi de suite. Dans ce type de modèle, la signification d'un concept est décrite par l'ensemble du réseau tel qu'il se ramifie à partir du noeud qui représente ce concept.

Illustration d'un réseau sémantique

Les liaisons entre concepts sont de différents types; les liaisons de superordination-hyponymie occupent une place particulière. Un premier arc partant d'un concept pointe vers une catégorie superordonnante, et les arcs suivants pointent vers des propriétés qui spécifient cette catégorie superordonnante.

Le canari se définira donc comme "un oiseau jaune qui peut chanter" : on a un arc vers un concept superordonnant ('oiseau) et deux propriétés spécifiantes (jaune, peut chanter). Bien entendu, "oiseau", "jaune" et "chanter*" sont eux-mêmes des concepts, à définir à leur tour par des propriétés, selon le même schéma.

Ceci permet de souligner une propriété fondamentale des structures sémantiques ainsi décrites: elles sont **hiérarchisées**. Conçues de cette manière, les structures sémantiques présentent l'avantage d'être particulièrement économiques, pourvu que les propriétés spécifiantes soient attachées le plus haut possible dans la structure. En effet, si la personne stocke en mémoire la proposition <Se mouvoir* (animal)>, elle est dispensée de stocker cette information au niveau des catégories inférieures. Si la personne

a stocké qu'un oiseau est un animal, et qu'un animal peut se mouvoir, elle peut en inférer qu'un oiseau peut se mouvoir.

Dans ces conditions toutefois, on voit que les structures sémantiques ne sont entièrement utilisables qu'en mettant en oeuvre des processus d'inférence. Si, par exemple, on demande à quelqu'un de dire si la phrase "Un canari peut voler" est vraie ou fausse, il ne trouvera pas cette information comme telle en mémoire; il faudra qu'il fasse une opération du type "Un canari est un oiseau; un oiseau peut voler; donc un canari peut voler". Par contre, "Un canari est jaune" devrait être vérifié plus rapidement, puisque cette information se trouve stockée comme telle en mémoire. Il faudra plus de temps encore pour vérifier "Un canari a une peau" puisque l'opération d'inférence est plus complexe: "un canari est un oiseau; un oiseau est un animal; un animal a une peau; donc un canari a une peau". Ces prédictions de temps de réponse ont été confirmées par plusieurs travaux inspirés du modèle en réseau.

Voir ci-dessous un fragment de réseau sémantique:

animal	avoir une peau se mouvoir
oiseau	avoir des plumes voler*
canari	jaune chanter*

Modèle de Collins et Quillian

Illustration de Temps de vérification, in JC Fig. 6, p. 33

Modèle en réseau bis: Spreading-activation model

Plusieurs difficultés (que nous n'avons pas le temps de détailler ici) ont mené Collins et coll. à proposer une version remaniée du modèle en réseau sous le nom de "spreading-activation theory of semantic processing". Ce modèle postule que le réseau est organisé sur la base de la similarité sémantique entre les concepts: plus nombreuses sont les propriétés que deux concepts ont en commun, plus nombreux seront les arcs reliant ces deux concepts par l'intermédiaire de leurs propriétés communes, et plus étroitement ces deux concepts seront liés.

Illustration des liens possibles entre deux concepts

Voir page suivante

Quand un concept est activé, cette activation se ramifie dans le réseau environnant avec une intensité décroissante. On postule aussi que les arcs sont de longueur variable, c'est-à-dire qu'ils demandent des temps variables pour être parcourus; la distance, et donc le temps, ne dépend plus seulement du nombre de noeuds à franchir. Enfin, on renonce à l'idée que le stockage doive être parfaitement économique.

Propagation d'une activation: Spreading activation theory

Les noeuds conceptuels d'un réseau sémantique peuvent être considérés comme changeant d'état dans le temps. Ils sont soit dans un état de repos, non activé, soit dans un état activé; dans ce second cas, leur niveau d'activation peut-être plus ou moins élevé. En outre cette activation, présente dans un noeud à un moment donné, est susceptible d'activer un noeud voisin, c'est-à-dire de se diffuser d'abord jusqu'aux noeuds les plus proches, puis, de proche en proche, aux autres noeuds voisins de ceux-ci. Cette diffusion de l'activation - on parle aujourd'hui de "propagation"- s'effectue entre deux noeuds en fonction de la force de la liaison qui les relie.

Le schéma théorique général est donc celui d'un réseau sémantique servant de support organisé à de l'activation et à sa propagation. Ce type de modèle comporte ou en tous cas permet une analogie forte avec une approche neurobiologique. L'activation est en effet un phénomène neuronal bien établi: un neurone peut être actif ou au repos, et la neurobiologie est en mesure de nous dire en détail en quoi consistent les phénomènes électriques et chimiques qui constituent cette activité de neurones isolés. Nous savons par ailleurs que les neurones sont interconnectés entre eux, de façon extrêmement complexe, dans une structure en réseau matériel reposant sur l'existence de liaisons synaptiques; et nous savons que l'activation circule à l'intérieur de ce réseau cérébral. Il ne s'agit toutefois qu'une analogie!

Modèle en traits

En réalité, les mêmes données peuvent être expliquées d'une façon différente: en supposant que les mots sont stockés séparément, chacun d'eux étant représenté par une liste d'attributs (une liste de "traits sémantiques". Dans ce cas, la vérification d'une phrase n'est plus la recherche dans un réseau, mais la comparaison entre deux ensembles de traits.

Un modèle de ce genre a été proposé par Smith et al. (1974), et étayé par de nombreuses données expérimentales. Dans ce modèle, chaque mot comporte un ensemble de traits, qui varient quant au degré avec lequel ils définissent l'appartenance à la catégorie. Par exemple, le trait "a des plumes" est essentiel au concept "oiseau"; d'autres traits (par ex. la taille, le fait de voler ou non, etc.) sont moins essentiels. La vérification de l'appartenance à la catégorie se ferait alors en deux temps: d'abord, une comparaison globale, aboutissant à une décision rapide si la similarité dépasse un certain critère (par exemple "Une hirondelle est un oiseau") ou au contraire est inférieure à un critère minimal (par exemple "Une hirondelle est un caillou"). Cette première comparaison revient en fait à évaluer la typicalité de l'exemplaire: la décision est très rapide pour une valeur très élevée ou très basse de typicalité. Entre ces deux valeurs, un second processus de comparaison serait mis en oeuvre, portant uniquement sur les attributs "définitoires".

La différence essentielle entre le modèle en réseau et le modèle en traits porte sur la part respective qu'ils attribuent au **stockage** préalable des informations et au **calcul** d'informations non stockées en mémoire.

Les relations catégorielles ("Un canari est un oiseau") sont pré-enregistrées pour Collins & Quilian; pour Smith et al., elles sont construites par un processus de comparaison. Par contre, pour certains attributs ("Un canari a un bec"), c'est l'inverse: calcul dans le modèle en réseau, pré-enregistrement dans le modèle en traits. En fait, il est bien évident que certaines informations n'ont jamais été enregistrées et doivent être inférées (par exemple "La Reine d'Angleterre a des poumons"); celles qui sont stockées l'ont été en fonction de ce que le sujet a appris (et pas en fonction de leur statut logique: par ex. "une baleine est un mammifère" a de fortes chances d'être une information apprise, et non pas le produit d'une comparaison entre les attributs de la baleine et ceux qui définissent les mammifères. A vrai dire, la question "stockage ou calcul" est une question dont la réponse variera selon les individus, et selon les informations concernées.

Cela étant, les deux types de modèles ne diffèrent que par leur formulation; en réalité, ils sont formellement équivalents. On peut sérieusement se demander si le paradigme expérimental très limité (la vérification de phrases simples) sur lequel ont porté la plupart des recherches

permet vraiment d'explorer de façon pertinente l'organisation et le fonctionnement de la mémoire sémantique. Plusieurs auteurs (comme Kintsch par ex.) considèrent qu'il serait plus instructif et plus fécond de faire porter les recherches sur un niveau supérieur à celui du mot.

(c) La représentation des connaissances

D'autres modèles sont plus vastes et plus ambitieux que ce que nous venons de voir. Ils concernent en effet la représentation des connaissances, dont la plupart sont acquises par le langage et sont par ailleurs formulables sous forme verbale. Ces modèles sont toutefois peu explicites sur les processus mêmes de compréhension du langage: ils s'attachent surtout à représenter la forme que prend, en mémoire, le résultat de ce processus.

Nous ne les examinerons pas ici. Nous y reviendrons toutefois plus loin lorsque nous parlerons de la compréhension de textes (voir modèle de Kintsch).

**Note de transition à propos des représentations
D'événements et d'actions.**

Les modèles en réseaux sémantiques ont d'abord cherché à décrire les concepts et relations entre concepts qui concernent les objets, les individus et les entités diverses réductibles au même schéma. Il est possible de ranger dans une même hyper-catégorie les représentations en mémoire d'êtres vivants de diverses sortes (y compris les humains), et les représentations des choses. C'est possible pour la simple raison que les unes et les autres de ces représentations partagent un certain nombre de propriétés cognitives fondamentales, dérivées de la perception, touchant l'espace et le temps, la permanence et le changement. (Par exemple, le fait d'occuper une partie de l'espace, avec une forme et un contour déterminés, le fait d'être dotées d'une constance perceptive, de pouvoir être touchées, de résister à la pression manuelle, de pouvoir être manipulées, modifiées et affectées par des actions diverses, etc.) Toutes ces représentations peuvent être regroupées dans une hyper-catégorie qu'on peut appeler la catégorie «objet». C'est essentiellement de cette catégorie de représentations que nous avons parlé ci-dessus. Pendant très longtemps, en effet, les recherches n'ont porté que sur les concepts qui sont exprimés par des noms et des adjectifs.

Actuellement, se développent des recherches nouvelles sur la sémantique des verbes, conçus comme porteurs des représentations d'événements et d'actions. Elles incluent l'étude des relations entre ces dernières sortes de

représentations et celles qui y "participent", celles de leurs agents et patients.

Ces recherches renouvellent notre conception même de ce qu'est un "concept"; elles nous forcent à prendre au sérieux l'idée qu'une signification de verbe est un concept, au même titre qu'une signification de nom ou d'adjectif, mais avec une structure conceptuelle différente. Les significations de verbe sont de contenu extrêmement varié. Ci-dessous, nous simplifierons considérablement les choses en ne nous intéressant qu'à deux sous-classes:

(a) Verbes désignant des événements (et particulièrement ceux qui dénotent un changement): par exemple "Kim a beaucoup maigri cet été".

(b) Verbes désignant des actions (verbes qui véhiculent l'idée d'une cause du changement; par exemple "casser", "remplir"). "Remplir" c'est causer un changement par lequel quelque chose qui était vide devient plein; "Casser" c'est causer un changement par lequel quelque chose qui était intact devient divisé en plusieurs morceaux.

Comment conceptualiser la structure sémantique de <événement>? On peut le faire en utilisant une forme logique comme <V (x)> où /V/ est le prédicat Verbe, et /x/ est l'ensemble des sujets (ou agents) possibles de V.

Soit par exemple <TOMBER (x)>

Signifie que le verbe "Tomber" a pour référence tous les événements dans lesquels quelqu'un ou quelque chose tombe . Cette extension est comparable à celle d'un nom comme "pomme", qui est l'ensemble de toutes les pommes possibles. La signification de "Tomber", pour un locuteur, est constitué par l'ensemble de toutes les représentations d'événements, chez ce locuteur, dans lesquelles quelqu'un ou quelque chose est perçu ou pensé comme tombant.

On peut présenter cette idée plus concrètement. Supposons que l'on vous demande:

1. Donnez moi des exemples de choses qui peuvent tomber;
2. Donnez-moi des exemples de choses qui ne peuvent pas tomber.

On pourrait aussi formuler ces questions autrement, en faisant référence à des catégories grammaticales:

1. Donnez-moi des exemples de noms qui peuvent être des sujets du verbe "tomber";
2. Donnez-moi des exemples de noms qui ne peuvent pas être des sujets de "tomber";

Vous n'auriez aucune peine à répondre à ces questions, même s'il est vrai que la démarcation entre 1. et 2. n'est pas pour vous parfaitement nette. L'expérience fait en outre apparaître des réponses catégorielles. En effet, "peuvent tomber" des catégories de choses usuelles (des chaises, des verres, des stylos), de gros objets (des arbres, des avions, des météorites), animaux

(des chevaux de course, des éléphants), humains (des skieurs, des alpinistes, des patineurs), etc. Tout le monde

Ce que cette procédure expérimentale rend bien manifeste, c'est que tout locuteur a dans son esprit une représentation de ce qui peut tomber ou ne pas tomber. C'est une forme de connaissance (implicite, mais sans doute universelle) d'une composante de la signification de "tomber".

Voir figure 7 in JFLN p. 273

Tout le monde a dans sa tête un ensemble représentatif des "tombants possibles", depuis les plus typiques jusqu'aux plus rares. La théorie dit que ces représentations font partie de la signification du verbe "tomber".

Recherches de McRae, Ferretti et al. (1997, 2001)

Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, 2001, 44, 516-547.

Thematic roles as verb-specific concepts. *Language and cognitive processes*, 1997, 12, 137-176.

Ils mesurent le temps de décision lexicale sur un nom (par ex. 'cambrioleur') présenté dans deux conditions:

(a) celle où ce nom est précédé par un verbe qui le prend normalement comme patient très typique: "arrêter - cambrioleur"

(b) celle où ce nom est précédé d'un verbe sans rapport avec lui, comme dans: "chanter - cambrioleur".

McRae et al. montrent que la décision est plus rapide dans la condition (a): c'est une démonstration expérimentale du fait qu'un verbe est capable d'amorcer la représentation d'un patient typique. Autrement dit, lorsqu'un verbe est présenté, non seulement il active dans l'esprit du locuteur sa propre signification, mais il active aussi (ou pré-active) la représentation de son ou ses patient(s) et (agent(s) les plus typique(s). L'idée qu'il existe une liaison *mentale* permanente entre la représentation des verbes et celle de leurs participants possibles (agent et patient) se trouve donc confirmée par les observations expérimentales.

Voir aussi JFLN, p. 329 - 330

Langage et L'Homme, 2002, 9-25 (Signification des verbes)

Une idée qui découle immédiatement de ce qui précède (et qui est très importante pour rendre compte de la compréhension du langage) est la suivante: pour former dans son esprit une phrase mentale particulière (par exemple que "la chaise est tombée") [qu'il s'agisse de la produire ou de la comprendre], une même opération est toujours nécessaire. Elle consiste dans tous les cas à "remplacer" le /x/ de <V(x)> par une représentation déterminée d'objet ou d'individu (dans notre exemple, remplacer /x/ par la représentation de /la chaise/, c'est-à-dire activer en mémoire la représentation de /la chaise/). La compréhension de la phrase "La chaise est tombée" implique donc 4 opérations majeures:

1. Activer la signification du mot "tomber"
2. Activer la signification de "chaise"
3. Juger (implicitement) que "chaise" fait bien partie des objets connus en mémoire comme susceptibles de tomber.
4. Assembler cognitivement les représentations de "chaise" et de "tomber".

Ce qui nous intéresse particulièrement ici, c'est le point 3. Pour comprendre la phrase "La chaise est tombée", il faut impérativement que, dans la mémoire sémantique à long terme du compreneur, la représentation de /chaise/ fasse partie de l'ensemble des représentations d'entités susceptibles de tomber. S'il n'en était pas ainsi, si la phrase comporte un mot qui renvoie à une entité ne faisant pas partie de cet ensemble (par exemple dans "La mer est tombée"), la phrase semblera à priori dépourvue de sens. Si le mot est seulement marginal par rapport à cet ensemble (par exemple dans "La décision est tombée"), la phrase sera comprise et considérée comme métaphorique.

Ce que nous avons montré, c'est qu'un esprit/cerveau identifie toujours un objet particulier (par exemple, une pomme) au moyen de l'opération cognitive "est un"; c'est-à-dire au moyen d'un processus de catégorisation, qui se produit aussi bien chez une personne qui perçoit directement la pomme que chez la personne à qui on parle d'une pomme. La pomme, perçue ou dont on parle, est toujours traitée par l'esprit/cerveau comme <un exemplaire de la catégorie des "pommés">. Pour la simple raison que cet esprit/cerveau possède dans sa mémoire sémantique une représentation catégorielle, le concept de <pomme>: ce concept est véritablement l'instrument du traitement cognitif de toutes les pommes particulières.

Cette façon de concevoir la catégorisation pour des catégories d'objets se généralise très bien aux catégories d'événements et d'actions; on peut donc dire qu'un esprit/cerveau identifie et catégorise un événement particulier de la même façon qu'il le fait d'une pomme particulière. Il forme de façon

immédiate et automatique une représentation de cet événement. Voir quelque chose qui tombe, c'est bien <voir un certain événement qui est celui de quelque chose qui tombe>, et qui se distingue de <voir quelque chose qui heurte quelque chose>. Voir quelque chose qui tombe, c'est insérer l'information qui entre par les yeux dans l'information mémorielle qu'on a dans la tête, celle où se trouve stocké tout le savoir qu'on a de ce que c'est que <tomber>.

Le petit enfant apprend cela très tôt; il dit "A fait poum" ('pour dire "est tombé"). Il apprend peu à peu que toutes ces représentations se disent, en français, "tomber", ou autrement dans une autre langue. Mais l'acquisition du langage par l'enfant est, avant celle des mots, celle des représentations qui les supportent. Elle repose d'abord, pour les verbes familiers de la langue, sur un apprentissage abstraitif en rapport avec la perception, et seulement en second lieu sur l'apprentissage des dénominations.

Ce qui est central, c'est que l'opération de catégorisation (c'est-à-dire l'opération de jugement implicite de type "est-un") est aussi importante pour les événements qu'elle l'est pour les choses et individus. L'esprit/cerveau catégorise les événements exactement de la même façon qu'il catégorise les objets et les individus, en les rapportant à des représentations conceptuelles présentes dans sa mémoire sémantique. Pour <tomber> comme pour <pomme>, cela se fait toujours par l'appariement d'un petit paquet d'information entrante (perceptive ou verbale) à une représentation catégorielle, ou conceptuelle.

Esquisse d'un grand réseau sémantique comme espace multidimensionnel dans lequel toutes les représentations de tous les mots appartenant au lexique mental d'un locuteur sont présentes; certaines de ces représentations sont reliées entre elles, par des liens de divers types (lien du type "est-un", ou du type "peut avoir comme patient", ou "peut avoir comme agent") assortis d'une valeur qui exprime le degré d'associabilité entre deux représentations. Voir JFLN 290 seq.

4. Du mot à la phrase

4.1 Introduction: la créativité linguistique

Contrairement aux systèmes de communication animaux (qui se limitent à un petit nombre de messages différents), le langage humain permet un nombre indéfini de messages possibles, par la combinaison d'un nombre fini de mots (eux-mêmes construits à partir d'un petit nombre de phonèmes). Cette propriété est ce qu'on appelle la "créativité" ou "productivité" linguistique.

La combinaison de mots produit une unité de communication qu'on appelle "phrase". Cette unité de communication élémentaire:

- + résulte d'une combinaison de mots soumise à certaines règles (on ne combine pas les mots n'importe comment!);
- + exprime une certaine signification, grâce aux relations entre les mots ("Michel embrasse Kim" ne signifie pas la même chose que "Kim embrasse Julie");
- + remplit une fonction communicative, c'est-à-dire permet d'accomplir un "acte de langage". Il ne s'agit pas seulement de coder verbalement un certain état de choses; il s'agit aussi de dire quelque chose dans un certain but (faire une affirmation, poser une question, donner un ordre, etc.);

Dans toute langue, la combinaison des mots obéit à un ensemble de règles qui constituent la syntaxe de cette langue. Tout locuteur connaît (implicitement) la syntaxe de la langue qu'il parle: il est capable de produire et de comprendre des phrases que, pour la plupart, il n'a jamais produites ni entendues auparavant; il est capable de reconnaître (très facilement et très rapidement) qu'une phrase est correctement formée ou non, etc. En d'autres termes, il possède une compétence linguistique.

Un premier problème (auquel les linguistes et les psychologues ont consacré beaucoup de temps) est de déterminer la nature de cette compétence, et sur quoi elle se fonde. Les règles sont-elles totalement arbitraires, et acquises par apprentissage? S'agit-il de structures innées relevant du bagage génétique de l'espèce humaine? Ou s'agit-il de structures fonctionnellement motivées, c'est-à-dire motivées par les nécessités fonctionnelles de la communication; on peut imaginer par exemple, qu'au cours de l'évolution, différentes "façons de dire" seraient apparues comme autant de solutions à des problèmes de communication: dire les choses de telle façon qu'on

réussisse à faire ce que l'on veut "faire" en parlant. (voir illustrations de ces deux conceptions au cours).

Un second problème est d'expliquer comment cette compétence est mise en oeuvre dans la production ou la compréhension de phrases. Autrement dit, il s'agit d'expliquer la performance: par quelles procédures (quels processus) est construite (en production) ou découverte (en compréhension) la structure syntaxique d'une phrase. Ces procédures relèvent-elles d'un "composant autonome" (y-a-t-il un module spécialisé dans le traitement de l'information syntaxique), ou ces procédures utilisent-elles d'autres informations (par ex. informations lexicales, sémantiques, pragmatiques)?

En résumé: la phrase a un sens qui ne se réduit pas à la somme des significations de mots qui la composent; le sens de la phrase constitue une organisation mentale spécifique. Comment se construit ce "sens de la phrase"? En quoi consiste-t-il? Comment est-il représenté dans l'esprit des interlocuteurs?

Les principales opérations nécessaires à la compréhension d'une phrase sont donc les suivantes (sans préjuger de l'ordre dans lequel elles s'effectuent):

Soit la phrase: "*L'espion a vu le policier avec des jumelles*"

Accès lexical

<Espion>: [personne chargée de recueillir des renseignements secrets]
[James Bond, Mata Hari, etc] [Nom masculin] [Singulier]

Analyse de la structure syntaxique

Segmentation et structure hiérarchique: analyse en constituants, représentée par figure ci-dessous

Analyse sémantique

VOIR (Agent: espion, Objet: policier, Instrument: jumelles)

Intégration avec le contexte

Qui est le policier?

Qui est l'espion?

Nous avons ci-dessus déjà examiné un certain nombre de questions relatives à l'accès lexical ; dans la suite, ci-dessous, nous examinerons les autres opérations.

4.2 L'analyseur syntaxique

4.2.1 Quand commence-t-on l'analyse syntaxique?

A priori, deux solutions sont possibles:

(a) Attendre: on attend d'avoir assez d'informations pour commencer à construire la structure syntaxique; on garderait donc en mémoire un certain "paquet" d'information avant de commencer l'analyse syntaxique.

(1) *Michel a acheté des fleurs pour Julie*

L'auditeur pourrait hésiter entre avoir entendu "A acheté" ou "a jeté"; "Pour" est ambigu: (a) des fleurs que Michel veut offrir à Julie; ou (b) des fleurs que Julie lui avait demandé d'acheter. Si on attend, on devrait garder toutes ces hypothèses en mémoire, ce qui risque de poser des problèmes (Plus les ressources cognitives sont utilisées pour du stockage, moins il en reste pour effectuer des traitements).

(b) Ne pas attendre: on commence tout de suite, sans attendre; on prend des risques, mais cela réduit la charge cognitive. Si on se trompe (c'est-à-dire si on construit une structure inadéquate), on peut toujours revenir en arrière, et recommencer (c'est-à-dire construire une nouvelle structure).

(2) *The florist sent the flowers ...*

En vertu du principe d'immédiateté, il semble que l'on ait affaire à une phrase déclarative ("Le fleuriste a envoyé les fleurs ..." où "fleuriste" serait l'agent de l'action "envoyer"). En anglais, ce début de phrase pourrait toutefois continuer comme

(3) *... was pleased.*

A première vue, "The florist sent the flowers was pleased" semble agrammaticale ("Le fleuriste a envoyé des fleurs était heureux"). Et pourtant, c'est bien une phrase grammaticale qui comporte une phrase relative dont on a omis le pronom (ce qui est possible et même fréquent en anglais); elle est équivalente à:

(4) *The florist who was sent the flowers was very pleased*

(4) *Le fleuriste à qui on avait envoyé des fleurs était très content.*

Si on avait adopté la stratégie "Attendre", on n'aurait sans doute pas été surpris par la fin de la phrase (3); l'effet de surprise indique bien que l'on n'a pas attendu pour commencer à analyser syntaxiquement la phrase (2).

4.2.2 Sur quoi se base l'analyse syntaxique?

Pour procéder sans attendre à l'analyse syntaxique, le système cognitif a recours à diverses "stratégies".

(a) Stratégies "syntaxiques": il s'agit de stratégies consistant à utiliser des indices disponibles dans la phrase et qui permettent de décider très vite

quels sont les constituants qui doivent être construits. Soit la phrase (5) ci-dessous :

(5) *Le professeur endormait les étudiants*

<Le + nom>: utiliser les mots fonctionnels qui annoncent des constituants (un, sept, et, qui, avec, etc.); par ex. <Le + nom> (Le prof . . .): construire un Groupe nominal (qui, s'il se trouve au début de la phrase, est vraisemblablement le sujet du verbe qui va suivre).

<-ait>: ce suffixe permet de penser qu'il s'agit sans doute d'un verbe? (Dans le dictionnaire, on ne trouve que 20 mots qui se terminent en "-ait")

<endormait>: supposons que c'est une forme du verbe "endormir"; dans ce cas, on peut s'attendre à certains compléments (par exemple quelque chose de vivant).

(b) **Stratégies de traitement**: il s'agit de stratégies liées aux contraintes de notre système cognitif de traitement de l'information (et particulièrement, liées aux limitations de notre mémoire). Ces stratégies visent à réduire la charge pour le système cognitif, par exemple en construisant de préférence la structure la plus économique, au risque de se tromper. Par exemple "la stratégie de l'attachement minimal" pour minimiser le nombre de noeuds dans la structure, et donc pour minimiser la complexité de la structure.

Recherches de Frazier et Rayner

Soit la phrase (6) ci-dessous :

L'espion regardait le policier avec des jumelles

(a) lecture 1: < . . . regardait avec des jumelles>

(b) lecture 2: < . . . le policier qui avait des jumelles>

La stratégie de l'attachement minimal prédit que, lorsqu'on lit des phrases qui contiennent un syntagme prépositionnel (avec) qui peut être attaché au syntagme verbal (regardait) ou au syntagme nominal (policier), on choisit la première solution parce qu'elle permet de construire la structure la plus économique (moins de noeuds, et structure moins profonde).

Voir ci-dessous Figures (indicateurs syntagmatiques) pour des phrases de différents types, montrant que cette stratégie marche bien pour certaines phrases, mais ne marche pas pour d'autres.

Cette stratégie fonctionnera pour

(1) *L'espion regardait le policier avec des jumelles*

Mais conduira à une erreur pour

- (2) L'espion regardait le policier avec un revolver.

Observations:

Le temps de lecture pour phrases du type (1) est plus bref que pour des phrases du type (2). En outre, l'analyse des mouvements oculaires montre des régressions (des retours en arrière) pour des phrases du type (2) lorsque la personne lit le mot qui rend l'analyse structurale initiale caduque.

(c) **Stratégies sémantiques:** ces stratégies consistent à construire la représentation la plus probable, la plus sensée (sur base des connaissances générales).

Par exemple, on lit à des enfants de différents âges des phrases comme ci-dessous (1) à (4), et on leur demande de choisir parmi plusieurs images possibles celle qui correspond à la phrase qu'on vient de leur lire.

- (1) Le chat a chassé la souris
- (2) La souris a été chassée par le chat
- (3) Le chat a été chassé par la souris
- (4) La souris a chassé le chat

Vers 3 ou 4 ans, les enfants choisissent systématiquement l'image qui correspond au sens le plus habituel: le chat poursuit la souris. Plus âgé, l'enfant cesse de faire des erreurs, mais le temps pour choisir une image "moins probable" par rapport à ses connaissances du monde, est plus long que celui pour choisir une image compatible avec ses connaissances.

4.2.3 Y-a-t-il primauté de l'information syntaxique?

Certains auteurs considèrent que certains processus de traitement du langage sont effectués par des "modules" spécifiques et autonomes, et que ces traitements ne peuvent pas être affectés par des informations qui ne relèvent pas de ces modules. Cela voudrait dire que dans le cas de compréhension de phrases, il existerait un module syntaxique qui effectue des traitements de type syntaxique sans tenir aucun compte d'informations sémantiques ou contextuelles. On parle à ce propos de "autonomie de la syntaxe". Selon cette conception, il y aurait un premier traitement de la phrase visant à dégager sa structure en s'appuyant uniquement sur des indices syntaxiques; le traitement sémantique n'interviendrait que dans un second temps.

Les observations de Frazier et al. rapportées ci-dessus (sur l'attachement minimal) semblent en faveur de cette conception: Temps de lecture plus bref

pour "L'espion observait le policier avec des jumelles" que pour "L'espion observait le policier avec un révolver" + Mouvements des yeux en arrière plus fréquent pour la phrase (2).

D'autres auteurs ne pensent pas qu'il y ait un module syntaxique autonome; au contraire, ils soutiennent une conception du traitement selon laquelle des interactions sont possibles entre des informations de divers types (syntaxiques, sémantiques, pragmatiques). Ils se demandent si les effets observés sur le temps de lecture sont bien dus à des facteurs syntaxiques, ou à des attentes sémantiques de la part du compreneur. Ils soulignent que dans les phrases utilisées par Frazier et al. (1982), le contenu sémantique fait attendre un complément de verbe ("L'espion observait l'homme" invite à préciser l'instrument de l'observation); mais on peut construire des phrases qui produisent plutôt l'attente d'un complément du nom.

Recherches de Taraban et McClelland (1988)

Attente sémantique en faveur de l'attachement minimal (phrases dans lesquelles on s'attend normalement à rencontrer un complément du verbe)

- (1) Le garçon coupa la pizza avec un couteau
- (1bis) Le garçon coupa la pizza avec des tomates

Attente sémantique en défaveur de l'attachement minimal (phrases dans lesquelles on s'attend normalement à rencontrer un complément du nom)

- (2) Le garçon commanda une pizza avec enthousiasme
- (2bis) Le garçon commanda la pizza avec des tomates

Ils retrouvent les observations de Frazier et al. (temps de lecture) pour les phrases de type (1) et (1bis), mais ils obtiennent des résultats inverses pour les phrases (2) et (2bis). Autrement dit, (1) est lu plus vite que (1bis); par contre, (2bis) est lu plus vite que (2). La stratégie d'attachement minimal ne semble donc pas systématiquement appliquée.

Recherches de Britt et al. (1992)

Donnent à lire des phrases du type

- (1) The boy hit the girl with the boomerang
Le garçon a frappé la fille avec le boomerang

Normalement les lecteurs interprètent cette phrase conformément au principe d'attachement minimal ("le garçon frappe avec un boomerang"); mais si on fait précéder cette phrase d'un contexte où il est question de plusieurs filles, dont une tient un boomerang en main, les lecteurs ne font pas cet attachement minimal, et interprètent la phrase comme "le garçon frappe la fille qui a un boomerang").

Plusieurs autres observations conduisent à penser que des informations lexicales et contextuelles peuvent interagir avec le fonctionnement du

processeur syntaxique. Certains auteurs vont même jusqu'à refuser l'idée même de processeurs distincts.

Recherches de Tyler et Marslen-Wilson

Ils font entendre à des sujets des phrases comportant une ambiguïté syntaxique, que le contexte antérieur permet de résoudre; par exemple:

- (1) If you walk too near the run way, landing planes . . .
- (2) If you have been trained as a pilot, landing planes . . .
- (1) Si vous passez trop près de la piste, les avions qui atterrissent . . .
- (2) Si vous avez été entraîné à piloter, faire atterrir les avions . . .

Immédiatement après le dernier mot (planes), on présente visuellement <is> ou <are> en demandant aux sujets de le prononcer aussi vite que possible. Le temps de réponse est nettement plus court pour <are> avec la phrase (1), et nettement plus court pour <is> avec la phrase (2). La structure grammaticale du syntagme ambigu (landing planes) est donc déterminée dès l'audition de ce syntagme, et sur des indices qui ne peuvent pas être syntaxiques, mais qui relèvent de l'interprétation que le sujet a élaborée de la phrase.

4.3 Syntaxe ou sémantique?

Des auteurs comme Marslen-Wilson et Tyler (1987) non seulement pensent qu'il n'y a pas de "composant syntaxique autonome", mais plus généralement critiquent l'idée de "niveaux de traitement distincts". L'activité du sujet consiste essentiellement à construire une interprétation de la phrase; ils pensent que cette interprétation commence dès le début de la phrase en s'appuyant simultanément sur tous les types d'information (lexicale, syntaxique, contextuelle) disponible.

En effet, que cherche l'auditeur lorsqu'il entend une phrase? Il cherche à comprendre ce qu'on lui dit. C'est ce travail d'interprétation sémantique qui est essentiel, et les indices syntaxiques (ordre des mots, catégories grammaticales, marques morphologiques) ne servent que d'indications pour cette interprétation. Il suffit d'utiliser ces indications, au fur et à mesure qu'elles se présentent, pour construire la signification; inutile de construire, en plus, une structure syntaxique. Comme dit Johnson-Laird, "Si vous savez lire les panneaux indicateurs, vous n'avez pas besoin de carte".

Dans cette perspective, ce qui caractérise les travaux récents est l'attention croissante que l'on porte au rôle du lexique dans le traitement des phrases. Ce rôle est de plus en plus pris en compte dans les théories linguistiques elles-mêmes, et est mis en évidence dans de nombreux travaux psychologiques.

Rôle du lexique

Recherche de Trueswell et al. (1994)

Les sujets doivent simplement lire des phrases qu'on leur présente sur écran; on enregistre leur mouvements oculaires. On leur fait par exemple lire des phrases telles que (1) et (2) ci-dessous:

- (1) The defendant examined by the lawyer turned out to be unreliable
- (2) The evidence examined by the lawyer turned out to be unreliable

Le début de la phrase (1) peut être interprété comme (1a) "Le témoin a examiné ..." ou comme (1b) "Le témoin qui a été examiné ..."; "témoin" est en effet aussi bien quelqu'un qui peut examiner quelque chose, que quelqu'un que l'on examine. Si le lecteur choisit (1a), il rencontre un problème dès qu'il lit la suite de la phrase.

Le début de la phrase (2) ne peut être interprété que d'une seule façon: "La preuve qui a été examinée..."; en effet, "preuve" ne peut pas être l'agent de l'action "examiner".

Les résultats montrent qu'il y a beaucoup de régressions oculaires pour des phrases du type (1), et pas de régressions pour des phrases du type (2). Ces résultats indiquent bien que les lecteurs ont d'emblée utilisé une information lexicale ("preuve" ne peut pas être agent) pour construire l'interprétation de la phrase.

Tous ces travaux invitent à considérer qu'on peut difficilement séparer le traitement syntaxique du traitement sémantique.

4.4 Représentation sémantique de la phrase

Le traitement de la phrase aboutit à la construction d'une représentation sémantique. Cette représentation est de nature abstraite, conceptuelle; elle est "amodale", c'est-à-dire indépendante de la modalité -verbale ou imagée- sous laquelle peuvent être matérialisées et transmises les informations.

4.4.1 Représentation amodale: ni verbale ni imagée.

L'indépendance du "sens de la phrase" par rapport à son support (verbal ou imagé) a été démontrée par plusieurs recherches expérimentales.

Recherche de Potter et al. (1987).

On fait entendre aux sujets des phrases (par exemple: "La jungle retentissait des cris d'oiseaux exotiques") suivies immédiatement de la présentation visuelle d'un mot ou d'un dessin (par exemple le mot "singe", ou un dessin représentant un singe). On demande simplement aux sujets de décider le plus vite possible si le mot ou le dessin s'accorde ou non avec le sens de la phrase.

Si ce sens de la phrase était représenté en mémoire dans un "format" spécifique (verbal ou imagé), on devrait observer un temps de réponse plus long dans l'un des deux cas (mot ou dessin) qui réclamerait une "traduction" d'une modalité à l'autre. Or aucune différence n'apparaît: la compréhension de la phrase se fait donc sous forme d'une représentation qui n'est ni verbale, ni imagée, mais qui peut indifféremment se traduire dans l'une ou l'autre des deux modalités.

L'indépendance du sens de la phrase par rapport à son support verbal peut être constatée facilement: lorsqu'on écoute une conversation, ou un discours suivi, on peut en général répéter mot pour mot la dernière phrase que l'on vient d'entendre; par contre, de ce qui précède (ce qui a été dit avant cette dernière phrase), on ne conserve que le sens général.

Recherche de J. Sachs (1967)

A montré que la formulation littérale d'une phrase était très rapidement oubliée, mais que sa signification était assez bien conservée.

Les sujets écoutent de petits paragraphes, tels que par ex. une brève histoire sur l'histoire de la découverte du télescope (voir exemple ci-dessous); dans chaque paragraphe, il y a une phrase-cible (indiquée *en italique* dans l'exemple ci-dessous).

Exemple: texte sur l'invention du télescope.

Voici une histoire intéressante à propos du télescope. En Hollande, un homme qui s'appelait Lippershey était un fabricant de verres de lunettes. Un jour, ses enfants étaient en train de jouer avec des lentilles. Ils découvrirent que les choses paraissaient beaucoup plus proches si on les regardait à travers deux lentilles maintenues à quelque distance l'une de l'autre. Lippershey commença aussitôt à faire différents essais, et ses recherches sur les verres grossissants suscitèrent beaucoup d'intérêt. *Il envoya une lettre à ce propos à Galilée, le grand savant italien. Galilée, immédiatement, réalisa l'importance de cette découverte et ...*

L'audition de ce texte était interrompue à un moment donné, et le sujet recevait une phrase-test en ayant simplement pour consigne de dire si cette phrase-test correspond à une phrase entendue.

La phrase-test pouvait être (a) tout à fait identique à la phrase-cible, (b) une version active ou passive de la phrase-cible, ou (c) une modification sémantique de la phrase-cible. Exemples ci-dessous :

Phrase-cible :

Il envoya une lettre à ce propos à Galilée, le grand savant italien.

Phrase-test identique :

Il envoya une lettre à ce propos à Galilée, le grand savant italien.

Phrase-test: modification formelle :

Une lettre à ce propos fut envoyée à Galilée, le grand savant italien.

Phrase-test: modification sémantique :

Galilée, le grand savant italien., lui envoya une lettre à ce propos

La phrase test était présentée soit immédiatement après la phrase-cible, soit après 80 syllabes (environ 25 secs) ou 160 syllabes (environ 50 secs) constituant la suite du texte.

Quand la phrase-test était présentée immédiatement après la phrase-cible, elle était classée correctement comme identique ou différente dans près de 90% des cas. Après un délai, qu'il soit de 80 ou 160 syllabes, la reconnaissance d'une modification reste bonne s'il s'agit d'une modification sémantique; par contre, dans toutes les autres conditions, la performance ne dépasse pas le niveau du hasard. Les modifications portant sur le sens des phrases sont très bien détectées, même après 50 secondes; par contre les modifications touchant la forme de surface de la phrase ne sont bien identifiées que juste après l'avoir entendue.

Insérer ici figure Résultats de Sachs, from YB

4.4.2 Représentation propositionnelle

La représentation sémantique de la phrase ne présente donc pas de forme syntaxique particulière; elle se présente plus vraisemblablement sous une forme propositionnelle (proposition logique); il s'agit d'un format qui établit les relations entre un prédicat et divers arguments (qui fait quoi à qui quand et avec quoi).

Une phrase comme "Michel est tombé dans un magasin de vêtements" donnerait lieu à la représentation propositionnelle suivante:

p1 TOMBER (Michel, L p3)
p2 PASSE (p1)
p3 DANS (Michel, L magasin)
p4 VENDRE (L magasin, O vêtements)

Pour l'exemple ci-dessus, la représentation propositionnelle de la phrase serait quelque chose d'équivalent à :

p(1) ENVOYER (Agent: Lippershey; Objet: lettre; Destinataire: Galileo)
p(2) A PROPOS DE (Lettre; Découverte)

4.5 Au delà de la représentation sémantique: le modèle mental

De nombreux travaux montrent que nous ne nous limitons pas à construire une représentation propositionnelle des phrases que nous entendons ou que nous lisons; nous construisons aussi une représentation de ce qui est décrit par la phrase ou le texte: nous construisons un "modèle mental" ou un "modèle situationnel".

Recherche de Bransford et al.

On fait lire des phrases. Après, on présente une autre série de phrases: pour chacune, le sujet doit dire si il l'a déjà lue ou non.

Soit par exemple, dans la phase de lecture, les phrases suivantes:

- (1) Trois tortues flottaient sur un morceau de bois, et un poisson nageait en-dessous d'elles.
- (2) Trois tortues flottaient à côté d'un morceau de bois, et un poisson nageait en-dessous d'elles.

Supposons que, dans la phase de reconnaissance, on donne ensuite aux sujets les phrases suivantes:

- (1bis) Trois tortues flottaient sur un morceau de bois, et un poisson nageait en-dessous de lui.
- (2bis) Trois tortues flottaient à côté d'un morceau de bois, et un poisson nageait en-dessous de lui.

On observe beaucoup de fausses reconnaissances (= dire "je l'ai déjà entendue" alors qu'on ne l'a pas entendue) si on présente (1bis) après avoir présenté (1); par contre, il y a très peu de fausses reconnaissances si on présente (2bis) après avoir présenté (2). Le sujet a donc très vraisemblablement construit un "modèle mental" de la phrase initiale: (1) et (1bis) donnent lieu au même modèle mental -ce qui explique les nombreuses fausses reconnaissances-, mais (2) et (2bis) ne donnent pas lieu au même modèle mental (le poisson a changé de place), ce qui explique qu'on observe peu de fausses reconnaissances

4.6. Au-delà du sens littéral

Dans ce que nous avons dit jusqu'ici, il n'a été question que du sens littéral d'une phrase. Cependant, dans la vie de tous les jours, il arrive fréquemment que l'on produise des phrases pour signifier autre chose que ce qu'elles signifient littéralement. Autrement dit, le vrai sens n'est pas le sens littéral, et c'est ce sens "non littéral" ("l'autre sens") que nous comprenons. Voir quelques exemples ci-dessous:

- (1) Pouvez-vous me dire l'heure?
- (2) Certaines maisons sont de véritables tombeaux
- (3) On manque de place ici
- (4) On marche sur des oeufs
- (5) Les buildings sont les girafes de nos villes
- (6) Ce pianiste est un véritable boucher

Beaucoup de travaux se sont intéressés à la question de savoir comment on procède pour comprendre ces phrases qui sont tout à fait correctes, mais dont le sens littéral n'est pas pertinent dans le contexte où elles ont été prononcées. Schématiquement, cette compréhension se ferait en trois étapes:

- + Comprendre le sens littéral [1].
- + S'apercevoir que le sens littéral n'est pas pertinent [2].
- + Ré-interpréter l'énoncé et dériver le sens non-littéral [3].

Examinons de plus près l'étape [2], en nous demandant comment on peut décider que le sens littéral est non pertinent.

4.6.1. Le principe de coopération.

Grice (1979) explique que, dans certains énoncés, le sens littéral bafoue manifestement un principe général de la communication, de l'échange. Voyons rapidement en quoi consiste ce principe; nous verrons ensuite comment on peut le bafouer.

Le langage humain sert à communiquer, c'est-à-dire à transmettre de l'information avec l'intention d'affecter le destinataire. Pour que cela fonctionne, il est nécessaire que les interlocuteurs coopèrent; en réalité, on peut considérer que les interlocuteurs coopèrent "naturellement". Par exemple, ils parlent assez fort, ils ne parlent pas tous en même temps, ils

essaient de construire des phrases correctes, etc. Mais ils sont aussi coopératifs en ce qui concerne le contenu et la manière de l'exprimer (par exemple, ils tiennent compte de ce que leur interlocuteur sait déjà, de ce que leur interlocuteur croit, etc.). Normalement, les interlocuteurs respectent une règle générale de coopération : cela signifie qu'ils se comportent de façon à maximiser leur compréhension mutuelle. Ils ne disent pas n'importe quoi ; et ce qu'ils choisissent de dire, ils ne le disent pas n'importe comment.

Le philosophe Paul Grice (1975) fut un des premiers à vouloir expliciter ce principe général de coopération. Il a proposé de formuler ce principe général sous la forme de 4 règles spécifiques (Grice parle de "maximes conversationnelles") qu'il appelle la maxime de quantité, la maxime de qualité, la maxime de relation et la maxime de modalité. On peut préciser brièvement ces 4 règles de la manière suivante :

- *Quantité* : Donner autant d'information que nécessaire, ni plus, ni moins.
- *Qualité* : Ne dire que ce que l'on croit vrai ; ne pas affirmer de choses au sujet desquelles on manque d'informations.
- *Relation* : Donner des informations qui soient en rapport avec le thème traité ; parler à propos, être pertinent.
- *Modalité* : S'exprimer clairement, parler sans ambiguïté, être bref et ordonné.

On pourrait certes penser que ces principes de coopération sont d'une très grande trivialité (ce qui explique sans doute que plusieurs auteurs ne se sont pas privés de les compléter, en considérant peut-être que plus compliqué ferait aussi plus sérieux). L'intérêt de tels principes, toutefois, devient plus manifeste quand on les examine dans leur fonctionnement ; ce sont ces principes en effet qui autorisent l'auditeur à faire certaines inférences ("invited inferences") quant aux intentions qu'avait le locuteur. On ne s'étonnera pas, par exemple, qu'un écolier puisse éprouver quelque difficulté à résoudre un problème formulé comme suit :

Un train parcourt la distance Bruxelles-Paris en 2h30. Calculez sa vitesse horaire moyenne, sachant qu'il parcourt 300 km et qu'il transporte 600 voyageurs.

Cet écolier pourrait en effet se demander ce qu'il y a lieu de faire de la dernière information, considérant, à juste titre au regard des principes de Grice, que si elle figure dans l'énoncé du problème, c'est qu'elle est pertinente. Grice lui-même attire l'attention sur l'intérêt qu'il peut y avoir à commettre des infractions volontaires.

Grice cite l'exemple de ce professeur d'université américain, invité à rédiger une lettre de recommandation pour un de ses anciens élèves postulant un poste de prof. en philosophie, et qui la formule en ces termes :

Monsieur X tape remarquablement à la machine, et il a été assidu à mes cours.

L'auteur de cette recommandation (appelons le John Biggs), en disant nettement moins que ce qu'on attend de lui, viole le principe de quantité d'une manière si manifeste que le destinataire ne pourra manquer d'en conclure que Mr. X ne fera pas un bon prof. de philosophie.

Cet exemple montre combien l'intérêt des principes de Grice réside précisément dans les effets de leur violation potentielle, car le fait même de leur violation constitue une information. Voici le raisonnement suivi par celui qui lit cette phrase :

- (a) John Biggs a dit P.
- (b) Je n'ai aucune raison de penser que John Biggs n'observe pas le principe de coopération.
- (c) Mais pour cela, il faut que John Biggs pense Q.
- (d) John Biggs sait (et il sait que je sais qu'il sait) que je comprends qu'il pense Q.
- (e) John Biggs n'a rien fait pour m'empêcher de penser Q.
- (f) John Biggs veut donc que je pense Q

Cet exemple montre bien que le " sens pertinent " est le sens qui restaure le principe de coopération. La violation manifeste, évidente de ce principe est d'ailleurs à la source des actes indirects (e.g., " *Pouvez-vous me passer le sel ?* "). Dans un contexte où il est évident que la personne " peut ", il ne s'agit pas d'une véritable question (on ne pose normalement pas une question dont on connaît la réponse) ; il s'agit d'un " ordre " équivalent à " *Donnez moi le sel* ". Poser une question dont on connaît la réponse, c'est violer le principe de coopération ; considérer cette " question " comme équivalent à un " ordre ", c'est restaurer le principe.

4.6.1. Le terrain commun.

Ne donner ni trop, ni trop peu d'information, la donner à propos et clairement, cela est très bien ; cependant, c'est du point de vue du destinataire qu'il y a lieu de juger dans quelle mesure ces impératifs sont rencontrés. Comment le locuteur saura-t-il ce qu'il en est, sinon par des échanges constants entre les partenaires ? (Cfr en acquisition : enfants qui changent de réponse, et ce faisant passent d'une réponse correcte à une réponse incorrecte, quand l'expérimentateur leur posent plusieurs fois la même question). De tels échanges leur permettront d'acquérir une

connaissance mutuelle, et d'établir ce que Clark et Carlson (1981) ont appelé leur " *common ground* " (terrain commun), c'est-à-dire l'ensemble des connaissances, croyances et suppositions que les interlocuteurs partagent au moment où ils se parlent.

Il est important de noter le sens que l'on donne ici au terme " *partager* ". Pour qu'on puisse dire d'un locuteur L et d'un auditeur A qu'ils partagent une proposition p :

L et A partagent p

il n'est pas suffisant que :

L sait que p

A sait que p

Il faut encore que

L sait que A sait que p

A sait que L sait que p

Et même, il faudrait sans doute compléter cela par une série infinie de propositions du type:

L sait que A sait que L sait que p

A sait que L sait que A sait que p

Diverses recherches sur l'utilisation de la référence définie illustrent le rôle du terrain commun. Dans une de ces recherches (Clark, Schreuder et Buttrick, 1983), on montra à des étudiants américains une photo sur laquelle apparaissaient l'un à côté de l'autre le président Reagan et son conseiller Stockman. On avait établi au préalable que Reagan était bien connu de tous, et que Stockman était très peu connu. On posa aux étudiants une des deux questions suivantes :

Q.1 : *Vous savez qui est cet homme, n'est-ce pas ?*
You know who this man is, don't you ?

Q.2 : *Avez-vous une idée de qui est cet homme ?*
Do you have any idea of who this man is ?

A la question Q1, tous les sujets donnèrent des réponses du type " *Oui, c'est Reagan* " ; à la question Q2, 90% des étudiants pensèrent qu'on parlait de l'autre personnage. En fait, la Q1 fait croire qu'on présuppose que la personne à laquelle on s'adresse connaît la réponse, alors que c'est l'inverse pour la Q2, et c'est précisément cela qui a amené les étudiants à déterminer lequel des deux personnages était le référent de " *cet homme* ".

Dans une autre recherche, on montrait une photographie avec quatre fleurs, dont une était soit légèrement, soit nettement mise en évidence. On posait la question suivante :

Comment décririez-vous la couleur de cette fleur ?

Lorsqu'une des fleurs n'était que légèrement mise en évidence, la majorité des personnes interrogées répondaient : " laquelle ? ". Quand elle était nettement mise en évidence, la majorité des personnes donnaient d'emblée la couleur de cette fleur-là.

4.6.2. Aspects développementaux de la coopération.

4.6.2.1 Adaptation à l'interlocuteur

Traiter en détail de la façon dont les enfants " coopèrent " avec leur interlocuteur exigerait que l'on examine tous les aspects du développement de la compétence conversationnelle. Faute de temps, nous nous limiterons ici à aborder quelques-unes des questions relatives à cette compétence.

L'étude minutieuse des interactions adulte-enfant a montré que l'acquisition de certains principes du fonctionnement conversationnel est très précoce, et qu'elle se fait à la faveur d'interactions Mère-Enfant au cours desquelles la mère façonne le comportement de son enfant de telle manière qu'il acquière , par exemple, la maîtrise de la règle de l'alternance des rôles ou la maîtrise de certaines procédures d'établissement de référents communs (cfr. recherches sur l'évolution du pointage :grasping, reaching, pointing, labelling).

Un aspect essentiel de la compétence conversationnelle concerne la capacité de l'enfant à construire ses énoncés en fonction de son interlocuteur. De nombreux travaux ont porté spécifiquement sur le développement de cette capacité. Ces travaux montrent que dès 4 ans l'enfant modifie son langage en fonction de l'âge de son interlocuteur.

Shatz et Goldman (1973) ont demandé à des enfants de 4 ans d'expliquer le fonctionnement d'un jouet à un adulte, puis à un autre enfant de 2 ans. Brame-Mouling (1977) a demandé à des enfants de construire une histoire à partir d'une série d'images et de la raconter à des enfants de leur âge ou à des enfants plus jeunes. Il ressort de ces travaux que dès 4 ans, l'enfant utilise une variété étonnante de moyens visant à établir la communication avec des plus jeunes :

- + intonation plus marquée (prosodie accentuée)
- + pauses plus nombreuses
- + débit plus lent
- + remplacement d'un mot difficile par un plus simple
- + reprise du substantif plutôt qu'usage de pronom
- + fréquence accrue du discours direct
- + ajout d'explications complémentaires

Cette capacité d'adaptation a été confirmée par d'autres études portant sur des conversations spontanées d'enfants de 3 à 6 ans soit avec leur mère, soit avec des enfants de leur âge, ou avec des enfants plus jeunes.

Qu'est-ce qui détermine cette adaptation ? Autrement dit : qu'est-ce qui pousse un enfant à modifier son langage pour s'adapter à son interlocuteur ? Plusieurs études montrent que c'est le niveau linguistique de l'interlocuteur, et non pas l'âge ni l'apparence physique, qui est la variable pertinente. Masur

(1988) par exemple a montré que des enfants de 4 ans sont capables d'ajuster différenciellement leur langage selon qu'ils s'adressent à des enfants de 2 ans caractérisés comme très verbaux (MLU compris entre 1,8 et 4,0) ou peu verbaux (MLU compris entre 1,0 et 1,5). On a également montré des adaptations différenciées chez des enfants normaux de 5 ans interagissant avec des enfants retardés mentaux de même âge chronologique, mais d'un âge linguistique inférieur.

4.6.2.2 Communication référentielle

Nous terminerons en mentionnant les travaux relatifs au développement de la capacité à communiquer un référent (= dire le plus clairement possible de quoi on parle).

A la suite des travaux de Piaget, on pensait généralement que le langage des enfants de 3 à 6 ans ne répond pas encore à une véritable fonction interpersonnelle ; l'incapacité relative des jeunes enfants à communiquer a, dès lors, été presque exclusivement interprétée en termes d'égoïsme infantin. Des travaux plus récents soutiennent cependant que le concept d'égoïsme n'explique que très imparfaitement le développement de la compétence communicative. En effet, de nombreuses données (comme celles que nous venons d'ailleurs de rapporter) montrent qu'il y a adaptation à l'interlocuteur, même dans le cas d'enfants très jeunes (au stade " pré-opératoire "). L'hypothèse de l'égoïsme donne à penser que les messages du jeune enfant n'ont en quelque sorte qu'une signification " privée ", sans signification " publique " : l'enfant parlerait pour lui-même, se comprendrait lui-même, mais ne se soucierait pas de savoir si les autres le comprennent.

Asher et Oden (1976) rapportent des observations qui infirment cette idée. Dans cette étude, les enfants reçoivent une liste de couples de mots ; dans chaque couple, un mot est souligné (= le référent), l'autre mot ne l'est pas (= le non référent)

<u>Horloge</u>	Montre
Rivière	<u>Océan</u>
Enfant	<u>Bébé</u>
<u>Vélo</u>	Voiture
Etc.	

Pour chaque couple, l'enfant doit produire un message d'un seul mot devant permettre à une autre personne de choisir le référent parmi les deux mêmes mots (mais dont aucun n'est souligné).

Asher et Oden ont montré que les enfants émettant des messages ambigus pour un adulte (par ex. pour le couple " Horloge - Montre " : produire le message : " heure ", ou " aiguilles ") ne pouvaient pas mieux identifier le mot référent lorsqu'on leur représente la même liste avec, pour chaque couple, le message qu'ils ont eux-mêmes produit.

En réalité, pour " réussir " dans cette tâche de communication référentielle, il faut faire deux types d'opérations : (a) activer un certain nombre de réponses possibles [par ex. heure, temps, aiguilles, sonnerie, grand, etc.], et [b] pour chaque réponse possible, évaluer sa pertinence en la comparant au référent et au non-référent. La conclusion de ces recherches est que le jeune enfant, en situation de communication, ne fait pas à cette comparaison (alors qu'il en est tout à fait capable dans une tâche de discrimination perceptive : dire qu'un son C ressemble plus à un son A qu'à un son B). Whithehurst et Sonnenchein (1984) concluent que " the young child does not know that to communicate referentially is to describe differences " ; autrement dit, le jeune enfant ne sait pas qu'il faut décrire en quoi une chose est différente d'autres choses avec lesquelles l'interlocuteur pourrait la confondre

Illustrations : Figures du Tangram