Psycholinguistics – How We Process Language

I. Language Processing

A. **psycholinguistics** = the study of language processing mechanisms. Psycholinguistics like to study how word meaning, sentence meaning, and discourse meaning are computed and represented in the mind.

A. the burning question: "How is language done?"

A. Mostly an unconscious process. Example: We think when we're reading words on a page that it's a smooth process, but our eyes actually jerk across the page – a process called **saccadic** motion. (The individual jerks by our eyes are called **saccades**.)

II. Methods of Psycholinguistic Research

A. Naturally Occurring

1. **Spoonerisms** = "slips of the tongue".

Ex: The **d**ear old **q**ueen = The **q**ueer old **d**ean. (segment switching)

Ex: The **rules** of **word** formation = The **words** of **rule** formation. (word switching)

Ex: I'd forgot**ten** about that = I'd forgot about**ten** that. (morpheme switching)

What it shows: The entire phrase must be planned in advance, or else we couldn't switch segments, morphemes, and words like this. This reveals something about the manner in which sentence (and phrase) production is planned in the mind.

2. **tip-of-the-tongue** phenomenon: trying to access a word based on meaning, spelling, initial letter, rhyme, etc.

Ex: "Oh, it's that color that's really bright green….and it's also a really strong liquor…but it sounds a bit like "loose"….it starts with a "sh" sounds….chartreuse! That's it!"

What it shows: how words are organized in the mind = **mental lexicon**. Access of the mental lexicon must be very quick, since word recognition takes just 1/3 of a second.

Related Questions, regarding the mental lexicon:

1) How are entries in the mental lexicon linked to each other?
1) How are entries in the mental lexicon accessed?
1) What information is actually contained in an entry?

B. Experimental Techniques

1. **lexical decision**

a. Process: A word flashes on a computer screen, and the subject indicates whether the word is a real word or a nonsense word by pressing a button.

b. Things which are measured (**dependent variables**):

1. **response latency** = how long it takes the subject to decide if the word is a real word or a nonsense word
2. **response accuracy** = whether the subject is correct or not.

c. How this relates to the mental lexicon: in order to decide if a word is a real word or a nonsense word, the mental lexicon must be accessed.

1. real word: find the mental entry
1. nonsense word: realize that there is no mental entry

d. **frequency effects**: more frequent words take less time to access than less frequent words. This tells us that some part of the lexicon is organized by individual frequency of the word.

2. **priming paradigm**

a. Process: The word to be judged as real or nonsense (the **target** word) is preceded by a stimulus word (**prime** word)
Example: Judging "doctor" → precede with "nurse"

a. **Dependent** variable: if the prime affects the **response latency** or not.

a. **Priming effect**: Semantically related primes lead to faster response times. This tells us that the mental lexicon must also be organized by semantic relatedness, since a semantically related word preceding the target word makes the target word easier to access.

a. **Priming effect** also found for **orthographically related** primes, **phonologically related** primes (rhyming words), and **constituent morphemes**.

   a. orthographically related: *dock* primes for *doctor*
   a. phonologically related: *worse* primes for *nurse*
   a. constituent morphemes: *legal* primes for *illegality*

   These all show us ways in which the mental lexicon is organized – by spelling similarity, by phonological similarity, and by constituent morphemes. Thus, there are many different ways to prime for a single entry in the lexicon – suggesting that the entries are linked to each other in several different ways.

3. **Sentence Processing Techniques**

a. Basis: We have to **parse** a sentence to understand it → we process it by understanding the meaning and structure of its parts

a. **Timed-reading experiments**

   1. assumption: a difficult sentence takes longer to parse. Therefore, timing how long it takes to process the sentence allows us to rank how "difficult" different sentences are to process.

   2. **bar-pressing paradigm**: The subject reads a sentence one word at a time, and presses the space bar to indicate they have processed that word. One word appears on the screen at a time.

1. **moving-window paradigm**: A sentence appears with all the words dashed out and the subject presses the space bar to see the first word. Another space bar press turns that word to dashes and reveals the second word, and so on.

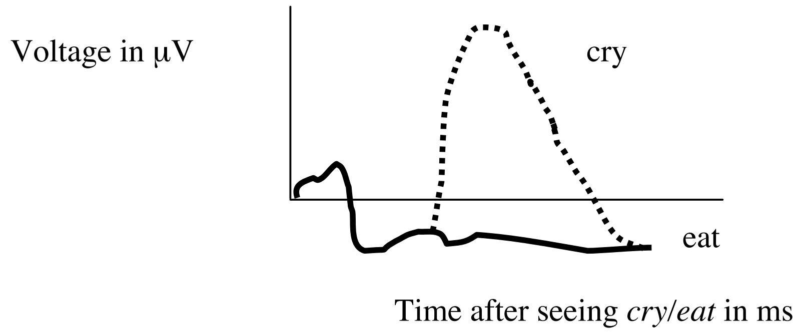
```
------------------------------------
Captain---------------------------
----------Hook--------------------
-----------------was--------------
-----------------------quite--------
------------------------------angry.
-------------------------------------
```

1. The pattern of how long it takes to process a word reflects the semantic and syntactic structure of the sentence.
   a. Content words take longer to process than function words. (semantic)
   a. Subjects pause at the end of clause boundaries. (syntactic)

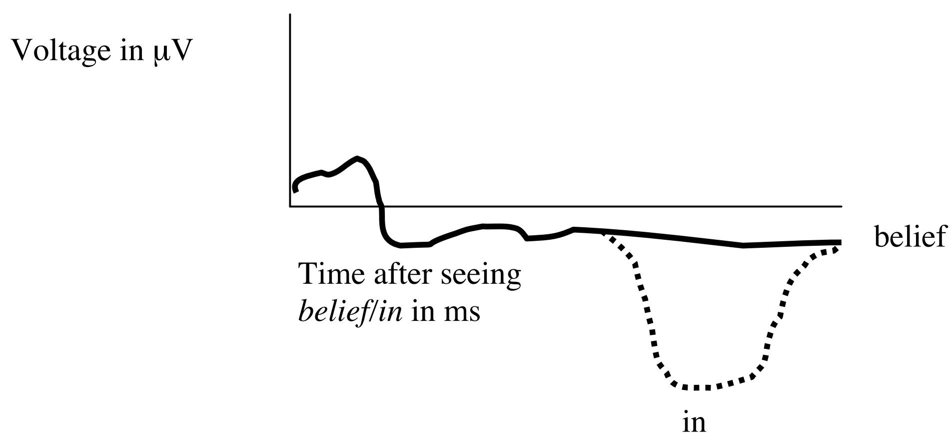
a. **Eye-movement experiments**: These experiments track the eye movements (**saccades**) of a subject while the subject is reading a sentence.
   1. Subjects tend to fixate on content words.
   2. Subjects' eyes move backwards in the sentence when a misparse occurs.
   3. Syntactically complex and semantically anomalous bits of sentences tend to create lots of backwards movements.
      a. syntactically complex: "*The defendant examined by the lawyer…*" = The defendant who was examined by the lawyer
      b. semantic anomalous: I like my coffee with cream and *socks*.

Assumption: The eye movements reflect processing. Thus, longer fixations and backwards movements reflect processing difficulty.

d. Brain activity: **ERPs**
   1. ERP = event-related potentials
   1. We can measure activity in the brain when a subject is reading a sentence.
   1. N400 = *negative* voltage change approximately *400ms* after a word is read which is semantically odd.
      Ex: "The pizza was too hot to *cry*." = N400 (instead of *eat*)



1. P600 = *positive* voltage change approximately *600ms* after a word is read which is syntactically odd.
   Ex: Sarah's belief in fairies vs. Sarah's *in* belief fairies



5. What this means: processing of sentences is immediate and "online" – happens as each word is read, rather than waiting until the end of a sentence/clause/phrase to put things together.

III. Linguistics and Language Processing

A. Top-down vs. bottom-up processing

1. **top-down**: We begin interpretation of a sentence spontaneously and automatically based on what information is available to us. For instance, we do not have to wait until we have analyzed all the phonemes in a sentence in order to understand it.
1. **bottom-up**: Do analysis to isolate phonemes, word boundaries, and relate these things to the mental lexicon. Can happen only piece by piece – no forward projection, no prediction.
1. Comparing **top-down** and **bottom-up** processing:
   In a lexical priming study, suppose a word is lexically ambiguous and so has two meanings. However, suppose that only one of those meanings is appropriate, given the syntactic structure of the sentence the word is in. A person using very strong top-

down processing would only be primed for the meaning which is appropriate, given the syntactic structure. A person using very strong bottom-up processing would be primed for both meanings, despite the fact that only one meaning is appropriate.

Ex: Hoggle fell gracelessly to the *ground*.
Top-down processing: prime only for *soil*
Bottom-up processing: prime for *soil* and *grind*

B. **Cohort Model** of word recognition

1. William Marslen-Wilson proposed that in word comprehension, words are recognized from beginning to end.
2. ex: hearing “crystal”. = [krɪstal]. First, we process the [k], and initially consider all the words that begin with [k]. All the words considered are called the **cohort**. Then we process the [r], and consider all the words that begin with [kr]. (The **cohort** is reduced from all the words that begin with [k] to all the words that begin with [kr]). And so on, until we process all the segments of “crystal”.
3. Shown with experiments that this is the case.
4. Suggests that the segment is a fundamental unit of auditory perception.

C. Evidence that syllables are also fundamental units of (auditory) perception

1. Syllables used successfully as primes in lexical decision tasks.
2. Word-blending tasks: subjects unconsciously split words at natural points in the syllable → the onset vs. the rhyme.
3. Example: “bark” + “meow” = “beow” (rather than “baow”)
4. Subjects prefer to create word blends according to the syllable structure of their language.

D. Morpheme Activation

1. Individual morphological components play a role in processing.
2. Individual morphemes in compound words automatically activated during word recognition.
3. Evidence: *crowbar* primes *bird*.

IV. Syntactic Processing

A. **syntactic parser**: (together with the grammar) guides the order of which elements of a sentence are processed and the manner in which syntactic structure is built up.

B. Some grammatically complex sentences are easy to parse and some grammatically easy sentences are hard to parse.

Ex: (complex sentence but easy parsing)
Sarah saw the goblin who displeased Jareth the other day.

Ex: (easy sentence but hard parsing)
The horse raced past the barn fell.

C. **garden-path** sentences: sentences which are easy for the grammar to produce, but hard for the parser. The parser is “led down the garden path” to the wrong structure.

Ex: The horse raced past the barn fell.
Parser’s garden path: “The horse raced past the barn” = simple sentence.
“The horse raced past the barn fell” = Ack!

Correct Parse: “The horse raced past the barn” = “The horse which was raced past the barn” → “raced past the barn” modifies “the horse”.
“The horse raced past the barn fell” = The horse (modifier) fell.

B. Garden path sentences reveal the preferences of the parser.
Ex: Given “The horse raced past the barn”, the parser can choose a simple Subj V PP structure, or a Subj + modifying phrase structure. The parsers chooses a simple Subj V PP structure because it prefers this analysis.

More examples of garden-path sentences:

1. The prime number few.
2. Fat people eat accumulates.
3. The cotton clothing is usually made of grows in Mississippi.
4. The man who hunts ducks out on weekends.
5. Mary gave the child the dog bit a bandaid.
6. The girl told the story cried.
7. I convinced her children are noisy.
8. Helen is expecting tomorrow to be a bad day.
9. I know the words to that song about the queen don't rhyme.
10. She told me a little white lie will come back to haunt me.
11. The dog that I had really loved bones.
12. The man who whistles tunes pianos.
13. The old man the boat.
14. The raft floated down the river sank.
15. We painted the wall with cracks.
16. The tycoon sold the offshore oil tracts for a lot of money wanted to kill JR.

C. The key: these sentences are ambiguous at the beginning. There are two structures the parser can choose up until the disambiguating word is reached. Sometimes context can help the parser decide which structure to choose.
Ex: There were these two horses. And they both were raced that day by an expert jockey. He raced one of them past the barn, and one of them past the house.

The horse raced past the barn fell.

Why does context help here?
We have two horses in the discourse. As soon as we mention *The horse*, we seem to assume that there is a unique horse we are talking about. Therefore, we are looking for some kind of description to attach to *The horse* in order to pick out which one we’re talking about. The parser would be pushed by **top-down processing** to interpret *raced past the barn* as a modifier to *the horse* so we can tell which one we’re talking about.

D. Parsing Strategies (in the absence of context)
1. minimal attachment: no extra syntactic nodes unless we have to.
Ex: “I convinced her children…”
→ **Subj V Obj** instead of **Subj V Obj [$_{CP}$ Subj…**
Garden Path: “I convinced her children are noisy.”
~~→ **Subj V Obj**~~ → **Subj V Obj [$_{CP}$ Subj V Adj]**

2. late closure: prefer to attach words to the node currently being processed.
Ex: "The old man…"
→ **Subj** instead of **Subj V…**
Garden Path: "The old man the boat."
→ ~~**Subj**~~ → **Subj V Obj**.

Exercises

1. Experimental Things
Muffin has been dabbling in psycholinguistic experimental techniques and has conducted an experiment to see if a priming effect was found for morphological roots on suffixed past tense forms in a lexical decision task.

a. Give an example of a morphological root for a suffixed past tense form.
b. What would be the prime in this experiment? What would be the target?
c. What is the dependent variable?

Suppose Muffin believes that there *is*, in fact, a priming effect found for morphological roots of suffixed past tense forms in a lexical decision task.

d. What sort of data would she be looking for? (How would this priming effect manifest itself?)

2. And More Experimental Things
Muffin has been dabbling with sentence processing experimentation. She has been testing pairs of sentences like the following:

i.The butter that had been melted in the pan started to burn.
ii. The butter melted in the pan started to burn.

a. What relation do these sentences have to each other? (Hint: Think sentence relations)
b. Muffin was using a moving-window paradigm. Show what the subject would have seen after 3 space bar presses on sentence ii. (Draw what was on the screen for press 1, press 2, and press 3.)
c. Muffin found that subjects seem to be much slower pressing the space bar after seeing *started* in sentence ii than in sentence i. Describe what this would mean, in terms of how difficult to process sentence ii is compared to sentence i.
d. What do you think would cause the slowdown at *started* in sentence ii?
e. If Muffin conducted an eye-movement experiment using these same sets of sentences (like i and ii above), what would you expect a subject's eyes to do upon reading *started* in sentence ii?

3. Yet More Experimental Things
Muffin just couldn't get enough, so she tried her claw at ERP experimentation. Using sentences like sentence i and ii in question 2 above, Muffin noticed a change in the ERP response of subjects at *started*. Would you expect this response to be an N400 or a P600 response? Why?

4. Cohort Model

According to the cohort model, how many phonemes of each of the following words would have to be processed before a hearer would be sure which word had been spoken? For each word, write the word in IPA. For each initial cluster up until the word would have no cohort left, give one example of a member of the cohort.

Ex: “prey” = /prej/
/p/ = pool
/pr/ = prefect
/pre/ = prelude
/prej/ = prayed
Need to get to end of word boundary.

1. giraffe
2. scram
3. splat
4. zebra

5. Garden Paths.

For each of the sentences below a) indicate what the ambiguous region of the sentence is, b) indicate what the disambiguating word/region is, c) indicate what the initial parse would look like up until the disambiguating word *using parentheses to group constituents*, d) indicate what the final parse would look like *using parentheses to group constituents*, and e) indicate what parsing strategy (or strategies) would lead to the garden path and why.

Ex: The old man the boat.
a) The old man
b) the (or possibly the end of the sentence, since *The old man the boat hit is okay*. is a fine sentence.)
c) (The old man)
d) (The old) (man) (the boat).
e) late closure – prefer to continue processing the NP, rather than start the VP; minimal attachment – prefer to not posit a VP when all the words could fit under an NP.

1. The man who whistles tunes pianos.
2. The girl told the truth cried.
3. She told me a little white lie would come back to haunt me.
4. We painted the walls with cracks.