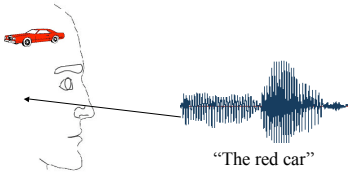
Introduction to Psycholinguistics

Lecture 8: Speech Production 1

Alissa Melinger

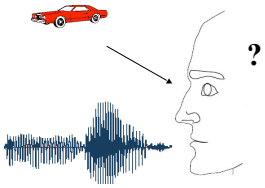
Computerlinguistik
Universität des Saarlandes

Translating Sound into Meaning



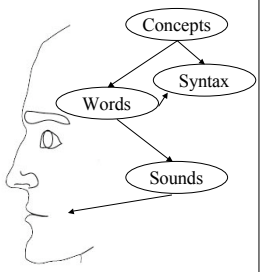
"The red car"

Translating Meaning into Sound



"The red car"

Speech Production: A Sketch

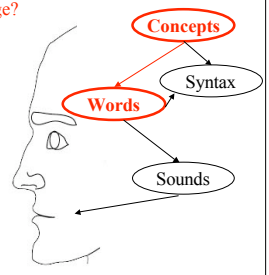


"THE RED CAR"

Issues on Speech Production

How do speakers choose the word(s) corresponding to the intended message?

Lexical Selection



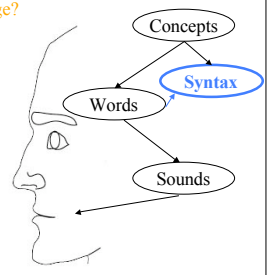
Issues on Speech Production

How do speakers choose the word(s) corresponding to the intended message?

Lexical Selection

How do speakers access a word's grammatical properties?

Grammatical Encoding



Issues on Speech Production

How do speakers choose the word(s) corresponding to the intended message?

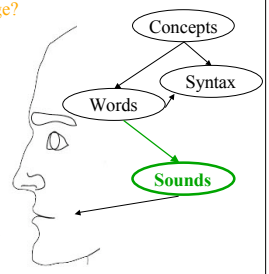
Lexical Selection

How do speakers access a word's grammatical properties?

Grammatical Encoding

How do speakers retrieve a word's phonological properties?

Morpho-Phonological Encoding



The miracle of speech production

- All humans acquire the language of their environment.
- By the age of about 5 or 6 we have a fully functioning linguistic system.
- As literate adults, we know between 50-100,000 words.
- We manage to produce between 2-4 words per second.
- We make only around 2 errors every 1000 words.
- Unlike in sentence processing or word recognition, we have nothing in the environment to help us. We create the output, the input comes from us.

Early Developments

- The early models of speech production (Fromkin, 1980; Dell, 1986) investigated the types of errors that occur in naturally occurring speech.
- Errors provide a window into the types of representations used in production and the types of computations involved.
- By observing the errors and forming generalizations over their patterns of occurrence, you can start to form hypotheses about how they happened.

Types of Speech Errors (Dell, 1986)

- Sound errors: Accidental interchanges of sounds between words.
  - Snow flurries ⇒ Flow snurries.
- Morpheme errors: Accidental interchanges of morphemes between words.
  - self-destruct instructions ⇒ self-instruct destructions
- Word errors: Accidental transpositions of words.
  - Writing a letter to my mother. ⇒ Writing a mother to my letter.
- Errors can have different forms
  - Anticipations: when a later element corrupts an earlier element
    - Reading list ⇒ leading list
  - Perseverations: When a later element is corrupted by an earlier element
    - hunting rabbits ⇒ hunting habits
  - Deletions: An output element is omitted
    - Same state ⇒ same sate.

Generalizations over the distribution of errors

- Units of all sizes can slip
  - Features, phonemes, segmental clusters, syllables, morphemes, words, phrases,
- When sounds exchange, they usually exchange with other elements in the same syllabic position.
  - Onsets with onsets; codas with codas.
  - Segments usually exchange within a phrase
- When words exchange, they usually come from the same grammatical class and have the same grammatical gender.
  - Nouns with nouns, verbs with verbs, masculine nouns with masculine nouns
  - Words usually exchange within a clause
- Errors involving sub-lexical elements don't respect this constraint.
  - Hunting rabbit --> hunting habit; rabbit and habit different part of speech
  - Suggests that sub- lexical units are stored/retrieved at a different level from words themselves. Grammatical information and phonological information are stored separately.
  - Suggests that lexical insertion processes are sensitive to grammatical information but phoneme and morpheme insertion are not.

More Generalizations

- When *apple* is exchanged for *banana*, the form of the indefinite article adjusts from an to a.
- Likewise, when a phonological error changes *A tin can → an in tan*, the indefinite article also adjusts for its new environment, suggesting that function words are inserted late.
- When nouns exchange, they leave their case-marked determiners behind
  - *I'm writing my letter a mother*
- When words exchange, they can strand their affixes
  - *Thinly sliced --> slicely thinned*
  - *Some swimmers sink --> Some sinkers swim*

Summarize what we know based on Errors

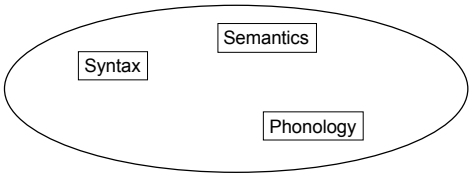
- Different computations operate over phonemes than operate over words.
  - Strong evidence for separating syntactic/semantic information and phonological/phonetic information.
- Grammatical information and phonological information stored and accessed separately.
- Function words inserted late into sentence frame

Another informative everyday phenomenon

- Tip-of-the-Tongue Phenomenon: we more or less know the word we want to say next, but are unable to bring it all the way to consciousness.
  - TOTs are more likely to arise on low-frequency words that have few close phonological neighbors.
  - We can often access information about part of speech or grammatical gender.
  - We can often access word length, stress pattern, # of syllables, or word onset information.
- More evidence for the separation of syntax/semantics from phonology!!!
  - We can have access to grammatical information without phonological access.
- Maybe we need a strong separation in the production system's architecture between syntactic information and phonological information.

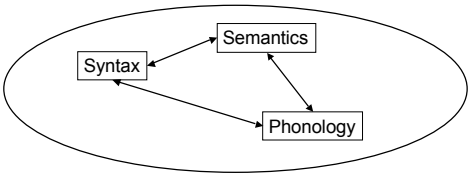
Speech Production Models

- There are several competing models of speech production, based on different types of primary data.
  - Reaction times, speech errors, patient data
- Speech production researchers agree on a few things.
  - We must distinguish semantic, syntactic and phonological types of information which are stored and accessed independently



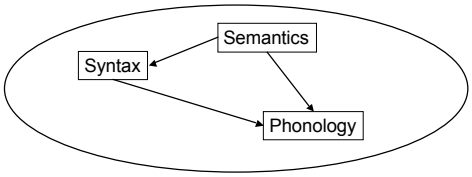
Speech Production Models

- How these respective stores of information are related to one another is a central question within the field.
  - Does information flow freely between all three components?



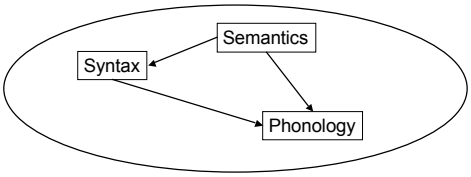
Speech Production Models

- How these respective stores of information are related to one another is a central question within the field.
  - Does information flow freely between all three components?
  - Or, does information flow only in one direction?



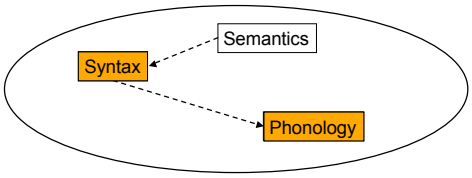
Speech Production Models

- How these respective stores of information are related to one another is a central question within the field.
  - Does information flow freely between all three components?
  - Or, does information flow only in one direction?
  - Are all the boxes linked? Or are some links not direct?



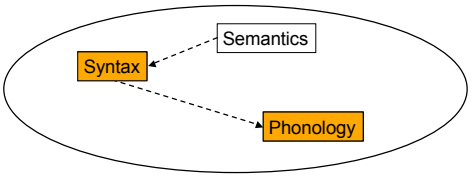
Speech Production Models

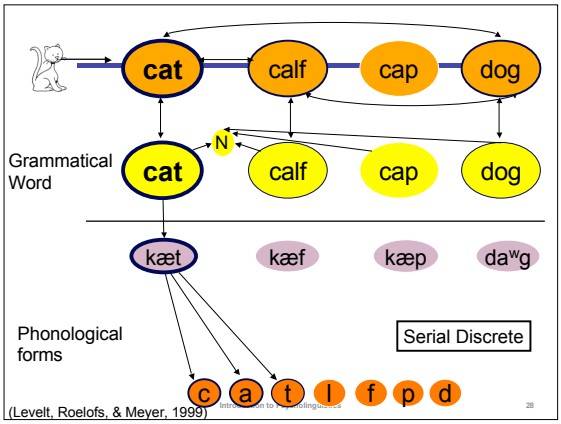
- How these respective stores of information are related to one another is a central question within the field.
  - Does information flow freely between all three components?
  - Or, does information flow only in one direction?
  - Are all the boxes linked? Or are some links not direct?
  - Does information flow continuously between the boxes?



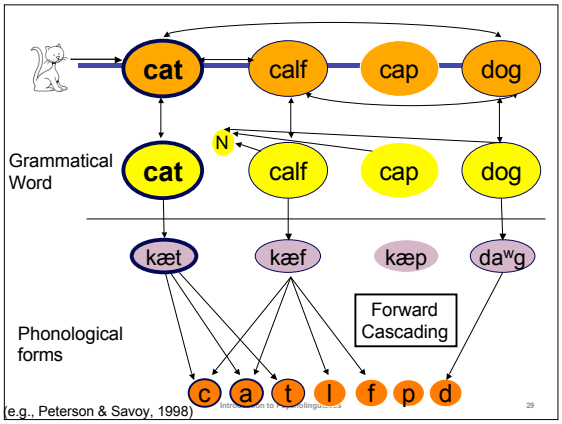
Speech Production Models

- How these respective stores of information are related to one another is a central question within the field.
  - Does information flow freely between all three components?
  - Or, does information flow only in one direction?
  - Are all the boxes linked? Or are some links not direct?
  - Does information flow continuously between the boxes? Or in discrete stages?

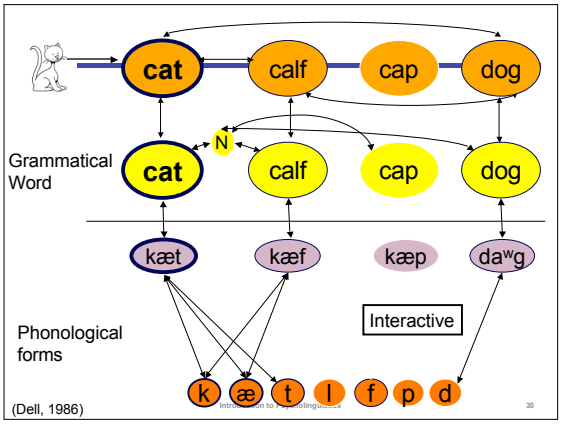




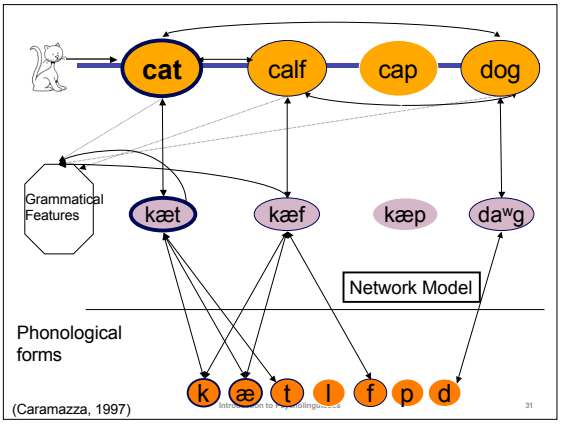
(Levelt, Roelofs, & Meyer, 1999)



(e.g., Peterson & Savoy, 1998)



(Dell, 1986)



(Caramazza, 1997)

Key Features of the Serial Discrete Model

- Wordforms are only activated after the lemma is selected; Only one wordform is activated
- Grammatical features must be selected prior to wordform encoding
- Lemmas compete for selection; No links between lemmas
- Effects at different levels shouldn't affect one another.

Key Features of Forward Cascading Models

- All active lemmas spread activation to their respective wordforms.
- Wordforms (also) compete for selection.
- Semantic and phonological effects are predicted to interact

Key Features of Interactive Models

- Activation is bi-directional. Wordform activation can influence lemma selection.
- Grammatical feature activation can affect lemma selection.
- Conceptual level = distributed feature network.
- Effects at different levels not isolated

Key features of network model

- Semantic representations componential.
- Semantic representations activate all lexemes in parallel that share semantic properties.
- Semantic representations *weakly* activate those syntactic features that have semantic reflex.
- Lexemes activate and allow selection of their associated syntactic features.
- Lexemes activate their associated segments and other form information.
- Activation is feedforwards and cascading.

Single Word Production

- We do make errors sometimes and we do have trouble finding our words. But, most of the time, we manage pretty well.
- How do we find the words to express our ideas?
- How do we manage with such speed and accuracy to find exactly that word which expresses best our meaning.
- We can be extremely eloquent sometimes, finding words that distinguish subtle shades of meaning to convey the exact nature of our thoughts.

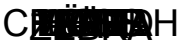
poodle
dog
animal
pet
puppy

white fluffy
tail fur

Name the pictures aloud in German as quickly as possible

Name the Pictures aloud in German as quickly as possible

Read the words aloud in German as quickly as possible



Observations

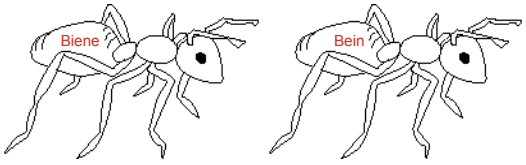
- Members of the same category interfere with each other.
- The more similar two concepts are to one another, the more interference that competitor will give you.
- CLAIM: When you have an idea in mind to express, activation spreads to all similar and related concepts and the corresponding words compete for selection.
- Words that compete for selection sometimes accidentally get chosen and we produce an error.

Lemma Competition

- All active lemmas compete for selection.
- The lemma with the highest activation level will get selected
- Once a lemma is selected, its grammatical features become available and are retrieved if needed in a given context.
  - Here you get information about the number, gender, or compatible phrase structures, etc. of the word, needed for articles, agreement, pronoun production, etc.
- The retrieval of these features drives phrasal construction and the insertion of function morphemes.

Glaser & Düngelhoff, 1984

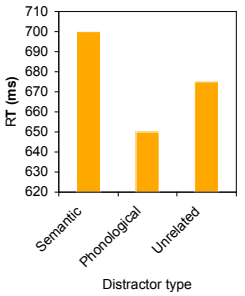
Picture-Word Interference



- Remember the Stroop Effect? BLUE PINK GREEN ORANGE
- A modification of this task can be used to investigate lexical selection.
- Provides strong evidence for competition between activated lemmas. Evidence against semantic competition too.

Picture Word Interference

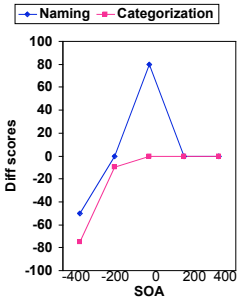
- Distractor words that are from the same semantic category as the picture name slow naming time compared to an unrelated distractor.
- Distractor words that are share phonological material with the target picture name speed response times relative to an unrelated distractor.
- Any distractor interferes relative to no distractor or a non-lexical distractor (XXX).

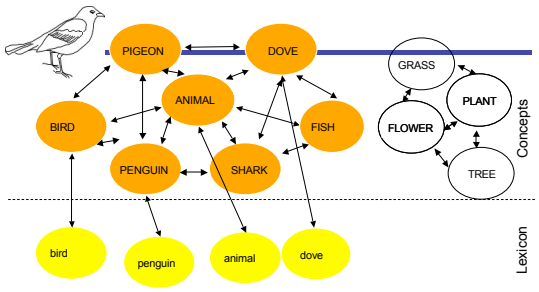


PWI: Stimulus Onset Asynchrony (SOA)

Timeline of effects for naming & categorization

- In Picture naming, semantic interference effects are observed at 0 SOA.
- In categorization, no interference effects observed.
- At long negative SOAs, a semantic facilitation effect is observed in both tasks.
- Since categorization, thought to be a primarily conceptual level task, does not reveal the same interference effects, this is taken as evidence for the interference effect being lexical rather than conceptual.





The lexical items compete for selection. Competition means, When one competitor becomes more active, target selection is delayed. Not a race to be most active.

Phonological Encoding

- The output of lexical selection serves as the input to phonological encoding.
- The selected lemma activates its wordform.
- Non-selected lemmas also activate their wordforms (Contra Levelt, 2001).
  - Much controversy on this point.
- For example, evidence that when the speaker wants to say 'horse' the lemma representation for 'goat' is also active (from semantic interference in PWI).
- Mixed evidence for whether the *wordform* for 'goat' is also activated.

Jescheniak & Schriefers, 1998

Activation of non-target wordforms?

- No effect of *goal*, which is phonologically related to the competitor *goat*.
- Soda is phonologically related to the non-target name *sofa*. The presence of this distractor *slows down* naming time, suggesting that the phonological representation of a non-selected lemma can be activated under certain circumstances.
  - Synonym = both names equally appropriate and leads to *double lemma selection*

Starreveld and La Heij, 1996

Activation of non-target wordforms?

- Mixed distractors do not behave like other semantic distractors.
- Mixed distractors do not behave like a combination of semantic and phonological distractors.
- There does seem to be some evidence for non-target wordform activation. But, the effects seem to be very weak and not always obtained.
  - In terms of the architecture of the production system, why might we want non-target word form activation? What role could it play in making us more effective speakers?

Phonological encoding

- The selected lemma and appropriate grammatical features then specify a morphological frame to be filled in by retrieved wordforms.
  - E.g., Stem + Affix + Affix
- The retrieved word forms activate, on the one hand, an ordered set of phonological segments and, on the other hand, a metrical frame.
  - Wordrform retrieval is frequency dependent.
    - High freq wordforms are retrieved faster than low freq wordforms.
- These two sources of phonological information must be assembled on-line; you don't retrieved assembled syllables.
- This assembly process unfolds from the beginning of the word to the end of the word.