VO EINFÜHRUNG IN DIE PSYCHOLINGUISTIK
ÁGNES LUKÁCS
WS 2011
LVA-Code: 160143

ZUSAMMENFASSUNG DER
POWERPOINT PRÄSENTATIONEN
1. INTRODUCTION ............................................................................................................................................4
1.1. Requirements for language (comprehension and production) ...................................................... 4
1.2. Topics in psycholinguistics ................................................................................................................. 4
1.3. Comprehension and production of language builds on several sub-tasks ..................................... 4
1.4. Generative Grammar .......................................................................................................................... 6
1.5. Psycholinguistic models ...................................................................................................................... 6
1.6. History of Psycholinguistics ............................................................................................................... 7
1.7. Cognitive science ................................................................................................................................. 8
1.8. Differences between human and nonhuman communication ............................................................ 9
1.9. Modularity (Fodor, 1983) ................................................................................................................... 9
1.10. Critical periods in language acquisition .......................................................................................... 10
1.11. Language universals ...................................................................................................................... 10
1.12. Psychological mechanisms ............................................................................................................... 11
1.13. Methods in Psycholinguistics ......................................................................................................... 13
2. BRAIN AND LANGUAGE: ......................................................................................................................15
2.1. Methods that utilize the anatomy of sensorimotor systems .............................................................. 15
2.2. Study of brain damaged patients (Paul Broca 1824 – 1880) .......................................................... 25
2.3. Measurement of brain activation during language tasks .................................................................... 30
2.4. Methods to locally disturb brain function ....................................................................................... 32
2.5. Neural components of different levels of processing ....................................................................... 33
3. LANGUAGE AND THOUGHT ................................................................................................................38
3.1. Relationship of cognition and language ............................................................................................ 38
3.2. Linguistic relativity ............................................................................................................................ 38
3.3. Language Evolution .......................................................................................................................... 43
4. LANGUAGE ACQUISITION ..................................................................................................................53
4.1. Stages .................................................................................................................................................. 53
4.2. Methods ............................................................................................................................................. 54
4.3. The acquisition of sounds ................................................................................................................ 56
4.4. One-word stage: lexical development ............................................................................................. 57
4.5. Vocabulary acquisition .................................................................................................................... 60
5. LANGUAGE ACQUISITION II ...............................................................................................................66
5.1. Grammar: two word stage ............................................................................................................... 66
5.2. Acquisition of Morphology ............................................................................................................. 68
5.3. How do we acquire language? ......................................................................................................... 70
6. VISUAL WORD RECOGNITION ..........................................................................................................75
6.1. Word recognition .............................................................................................................................. 75
6.2. Reading studies ............................................................................................................................... 77
6.3. Models of visual word recognition .................................................................................................. 79
6.4. Lexical ambiguity ............................................................................................................................. 86
7. MORPHOLOGICAL PROCESSING ......................................................................................................89
7.1. Decompositional or holistic? ............................................................................................................ 89
7.2. Word-level factors influencing processing ..................................................................................... 94
7.3. Words and rules ............................................................................................................................... 95
1. INTRODUCTION

1.1. Requirements for language (comprehension and production):

1. Biological equipment
2. Brain
3. A language with the right level of complexity
4. Abilities for acquisition
5. Knowledge of social context and rules
6. Knowledge about other’s knowledge and beliefs

Language is a communication system that is:
- rule driven
- symbolic
- hierarchically structured
- creative
- species-specific

Speech is not language! (sign languages!)

1.2. Topics in psycholinguistics:

- Sentence processing:
  How do we understand sentences.
- Morphological processing:
  How do we recognize morphologically complex words and how do we represent relations between related words.
- Spoken word recognition:
  How do we parse an acoustic stream into discrete units and then how do we recognize those discrete units as words of our language.
- Language Acquisition:
  How do we learn language; what are the stages of acquisition, what are the challenges the child faces.
- Speech Production:
  How do we find the words to express our thoughts; how do we assemble words into sentences; why do we make the errors that we do and not others.

1.3. Comprehension and production of language builds on several sub-tasks

1. The ability to produce and recognize sounds of the message
2. Recognition and production of words in the message, and the pairing of word forms with their meanings.
3. Processing and production of the grammatical structure of the message.
4. Evaluation of the message in the given context, or fitting the message to a given context.
1. The ability to produce and recognize sounds of the message
   • Relies on phonetic and phonological knowledge
   • Knowledge of the phoneme inventory of the language, the identification and distinction of different phonemes.
   • Successful use of a language also depends on knowledge of its phonotactics, or the ability to combine sounds into words
   • Which one is a better candidate for an English word?
     • blicket or lbnuxst

2. Word recognition and associating word forms with meanings
   • Builds on the lexicon: the mental inventory of words
   • Also requires knowledge of morphology, or the knowledge of word formation

Words:
Units of language
Crystal: the smallest unit of language that can form a complete utterance in itself, and that is separated by space from the rest of the text in writing.
Smaller units: syllable, sound, morphemes.
Tables, table – morpheme corresponding to the concept ‘table’,
-s: plural morpheme
Irregular forms (go-went or sing-sang versus walk-walked): do they still contain two morphemes?

The mental lexicon:
• Words are stored in the mental lexicon
• In many ways, similar to a real lexicon: pronunciation, meaning, written word form, category, syntactic roles, etc.
• But different principles of organization: e.g. frequency and meaning. Alphabetical order is not central.
• Ca 70,000 words in the adult lexicon (estimates vary between 15,000-150,000).

One of the central topics in psycholinguistics:
• The organization of the lexicon
• What is stored and how
• Lexical retrieval for production and comprehension
• How do we know whether it contains a word or not
• Are there differences between spoken and written word forms?

3. Processing and production of the grammatical structure of the message
   • The ability to combine words into phrases and phrases into well-formed sentences rely on syntax.
   • Knowledge of syntactic rules (grammar) generates well-formed (grammatical) sentences, and does not generate sentences that are not well-formed?

4. Evaluating the message in context
   • Successful expression and comprehension of a message also relies on pragmatic knowledge (formulating or decoding the message according to the given social context).
   • Knowledge of linguistic conventions that determine how we express ourselves in front of different audiences.
1.4. Generative Grammar

(Noam Chomsky 1950s, Standard Theory, inspired psycholinguistics)

The generative tradition:
- The aim of the study of syntax is to describe the set of rules that makes language production and comprehension possible.
- Distinction between idealized language competence and actual language performance
- Linguistics: competence--implicit knowledge of rules
- Psycholinguistic research: performance:
  - Why we sometimes make errors,
  - Why we make some errors not others,
  - Why we sometimes misunderstand some words or sentences,
  - How we learn a language,
  - How we accommodate speakers with thick accents,
  - How we deal with ambiguous words,
  - How we deal with ambiguous sentences
  - What makes a sentence ‘hard’ to understand?
- Grammar contains rewrite rules of phrase structure, which operate on terminal and nonterminal items
  S_NP + VP
  NP_Det + N
  VP_V+NP
  N_cat, dog, John, plate, etc.
  V_dropped, ate, hates, run etc..
  Det_a, the
  John dropped the plate.
- Standard Theory distinguishes between deep structure and surface structure.
- The deep structure of the sentence is the idea (together with its logical structure) behind it; the surface structure is the linguistic form which expresses the underlying thought.
- The surface structure is derived from the deep structure through transformations (special rewrite rules).
  Visiting relatives can be boring.
- Complex sentences are derived from several deep structures.
- The Derivational Theory of Complexity (DTC); Miller and McKean (1964)
  The vampire chases the ghost.
  The vampire does not chase the ghost.
  Does the vampire chase the ghost?
  Does the vampire not chase the ghost?

1.5. Psycholinguistic models

Realistic models of human sentence comprehension must account for:
- Language has structure
- Robustness to arbitrary input
- Accurate disambiguation
- Inference on basis of incomplete input
- Processing difficulty is differential and localized
Processing factors
• Besides abstract linguistic representations, other characteristics of words and sentences are also important
  – How often a word is used influences how easily we can retrieve that word.
  – How often a particular sentence type occurs in actual discourse influences how easily we can process it.
  – Practice does make perfect.
• Cognitive capacities relevant for language: memory resources, attentional resources, etc. These also affect linguistic performance in any task.
• Psycholinguistic research must take into account these factors
• Some factors that should be equated in the design of an experiment: word frequency, word length, concreteness, age of acquisition, part of speech, quality of picture, transitional probability of bigram, typicality or probability of X in Y context, morphological complexity, frequency of base, speech rate, duration of stimulus, number of words in a sentence, position on screen, repetition, ambiguity.

1.6. History of Psycholinguistics
• Started in the 1950s
  • First use of word: conference on Cornell, USA, 1951
  • Osgood and Sebeok, 1954: Psycholinguistics: A survey of theory and research problems
• Before that:
  – Sir Francis Galton examined word associations in 1867
  – In 1895 Meringer and Mayer analysed speech errors with surprisingly modern methods in Germany

Precursors
• Wundt: RI, first theory of speech production
• Broca and Wernicke: localization in the brain
• William and Clara Stern: diary data
• Skinner: behaviorism
• Chomsky: nativism

Modern psycholinguistics
• Early psycholinguistics theories based mainly on behaviorist principles.
  • 1959 Chomsky’s critical review of Skinner’s Verbal Behavior
• Beginning of 60s: psycholinguistics tries to relate language processing to transformational grammar
• It has left linguistics since, and became an autonomous field of study
• Roots both in psychology and in linguistics

Early psycholinguistics
• The language processor is a simple device that generates or accepts sentences by moving from one state to another.
• Two important influences: information theory and behaviorism.
  • Information theory
    – The role of redundancy and probability in language: at a give point, what is the most probable continuation of a sentence?
  • Behaviorism
    – stimulus, response, reinforcement to establish the relationship between them.
– Only relevant topic is behavior; language is one kind of behavior.
– Language acquisition also proceeds following the principles of reinforcement and conditioning

The influence of generative grammar
• Chomsky sweeping criticism: rapid change of views
• New kind of language theory: transformational grammar, which accounts for speakers knowledge of language besides the underlying structure of sentences.
• Part if psycholinguistics’ success is due to the fact that it tried to test Standard Theory and its consequences
• Not completely successful
  – It can tell us a lot about what we know when we know a language, and what constraints operate in language acquisition
  – But not about what mechanisms lie behind speech and language comprehension.

The 70’s
• Psycholinguistics becomes part of mainstream cognitive psychology
• Information processing or computational models
• Flowcharts (before translations into computer programs) with processing levels. How one level of linguistic representation is translated into another
• The mind uses rules that transform the input (speech or sight) into symbols. Cognition is symbol manipulation.
• The computational metaphor and experimental techniques made psycholinguistics a separate branch of science
• Just like in other fields of cognitive psychology, reaction time measurements became central.

1.7. Cognitive science
• A multidisciplinary approach: philosophy, linguistics, anthropology, neuropsychology and AI.
• AI: get the computer do something that requires intelligence
• Traditional AI research: analysis of human behavior, goals and plans
• Explicit
• A pro and con at the same time: loss of flexibility, many assumptions have to be built in
• Early optimism
  - Eliza (Wizenbaum, 1966)
  - Winograd: SHRDLU
• 1968-70 MIT
• Understands natural language

Important questions
• Is language species-specific?
• What is it that is innate and species specific in language?
• Is language domain-specific?
• Do we need special rules for language processing?
• Are processes or levels within language independent of each other? Do they work in cooperation?
• How are different processes of language related (e.g. reading and speech)?
• How do languages differ?
• What do we learn about language from acquired and developmental disorders?
• What are the subprocesses involved in language production and comprehension?
• How does the study of language relate to everyday language use?

1.8. Differences between human and nonhuman communication

• Animal communication restricted in its repertoire
• Animal language cannot be broken down into smaller units, it is not combinatorial
• When animals are taught language or use of a symbol system, the do not transfer this knowledge to their kids
• The relationship between a symbol and its meaning is arbitrary in human languages. This is much less so in animal communication.

1.9. Modularity (Fodor, 1983)

• A module is a specialized, encapsulated cognitive systems that has evolved to handle specific information types of enormous relevance to the species.
• Input → levels of processing → output
• What is the relationship between different levels of processing?
• Modules are
  – Encapsulated (it is impossible to interfere with the inner workings of a module)
  – Unconscious (it is difficult or impossible to think about or reflect upon the operations of a module)
  – Fast
  – Have shallow outputs (no information about the intervening steps that led to that output
  – Obligatory firing (operate reflexively, providing pre-determined outputs for predetermined inputs regardless of context)
  – Ontogenetically universal (develop in a characteristic sequence)
  – Localized (have dedicated neural systems)
  – Pathologically universal (break down in characteristic fashion following some insult to the system)
  – Domain specific (deal exclusively with a single information type)
  – It is assumed that learned systems do not display these all of these characteristics.

A modular view
Interactionism
• Modules are not encapsulated
• They cooperate
• Bottom-up and top-down processes

Modularity of language
• Big modularity:
  – How independent is language from the rest of cognition?
• Little modularity:
  – How independent are specific levels of language processing of each other?

1.10. Critical periods in language acquisition

Lenneberg (1967) critical period hypothesis:
– Certain biological events associated with language can only happen in an early critical period (hemispheric specialization).
– Certain language events have to happen in this period for language acquisition to follow a typical route
– Language is acquired most effectively in this critical period.

Evidence
• The two hemispheres are not fully lateralized at birth. An early left hemisphere damage or hemidecortication does not only result in permanent language impairment.
  – Maturational hypothesis: the two hemispheres of the brain are equipotential at birth.
  Language lateralization is strongest between 2-5 years, and closes by adolescence.
  – Invariance hypothesis: the innate organization of the left hemisphere makes it more prone or adapt to language (given up only for good reasons like brain damage to the left hemisphere)
• Second language learning.
• ‘Feral children’: Genie
  A weaker version of the critical period hypothesis:
  sensitive period, restricted to more complex aspects of syntactic processing

1.11. Language universals

• Substantive universals
  – Categories necessary for analysing or constructing languages: noun, vowel, subject, etc.
• Formal universals
  – Constraints on the types of linguistic rules
• Implicational universals: Of the form if x, then y
  – If a language marks gender on nouns, it will also mark it on pronouns.
  – If a language is predominantly VSO in its word order, then the adjective will most often follow the noun.

• Chomsky: there are certain universal constraints on rules and categories
• These constraints are biological and so innate
• Language acquisition device, Universal Grammar, that accounts for all possible human languages
• Principles and parameters
Does the language system use rules?
• linguistics: explicit rules
  – E.g. the plural form of nouns in English is formed by adding the –s morpheme to the stem
• For a long time, rules were proposed to operate in both comprehension and production.
  This view is more and more often challenged.
• E.g. in connectionist modelling
  – Builds on a plausible brain metaphor (?): processing is implemented by a densely
    connected network of neuronlike units. The model has to be very explicit.
  – Rethinking of linguistic representations. Connectionist models do not contain rules
    explicitly: these emerge as statistical generalisations over the data.

1.12. Psychological mechanisms
• serial and parallel processing
• bottom-up and top-down processing
• automatic and controlled processing

Serial vs. parallel processing

A serial model
Intention to convey an idea
Planning of clause structure
Retrieving lexical items
Retrieving phonological representations
A parallel model
Intention to convey an idea
Planning of clause structure
Retrieving lexical items
Retrieving phonological representations

Bottom-up and top-down processing
**Phonemic restoration effect** (Warren & Warren, 1970):

up  top  [peel, feel, wheel, heel]

bottom  down  */ee/l

Inability to report a disguised phoneme
Context can disambiguate it:
“It was found that the *eel was on the orange”
“It was found that the *eel was on the shoe” [peel and heel, respectively]

In general, how to make sense of speech in noisy rooms?
- Controlled processes: complex tasks that substantially draw on limited processing capacity
- Automatic processes do not tax limited resources
- Automatic: skilled word recognition
- Controlled: comprehension of main ideas

**Ambiguities**
- Local:
  - *Since Jay always jogs a mile seems like a short distance to him.*
  - *After the child had visited the doctor prescribed a course of injections.*
  - *The evidence examined by the lawyer turned out to be unreliable.*
- Persistent:
  - *The policeman saw the thief with the binoculars.*
  - *Visiting relatives can be boring.*
- Sources:
  1. Lexical ambiguity in meaning and/or in syntactic category (*The evidence examine...*).
  2. Ambiguous dependency (*The policeman saw the thief with the binoculars*).
1.13. Methods in Psycholinguistics:

Question is how do we get to know something about the processes taking place in the mind of a language user? Methods in psycholinguistics are by and large the methods of cognitive psychology. On the basis of experimental and observational data, researchers formulate models of what might be going on in our head.

**Experiments:**
- Most often used
- Online and offline methods
- Online – measures some kind of activity during language behavior:
  - Reaction times
  - Eye-movement registration
  - Speech monitoring
  - Brain imaging
- Off-line
  - Post-hoc task, memory involvement
  - Grammaticality judgments

**Experimental methods**
- Visual Comprehension (reading): lexical decision, naming, priming, subject-paced reading, eye movements, semantic categorization, brain imaging
- Auditory Comprehension (listening): lexical decision, priming, phoneme monitoring, brain imaging.
- Oral Production (speaking): picture naming paradigm, spontaneous errors, error inducing paradigm, brain imaging.
- Written Production (writing): thinking-aloud protocol, scriptlog

*Advantages and disadvantages of methods: unnatural settings, unnatural tasks, more natural tasks can be rather expensive.*

E.g., after 20 items per condition, 100 ms or 5% error difference in favour of short words => word length has an effect!

**Reaction times**
- How long does it take to read out a word?
- How long does it take to decide about a word?
  - Whether it is a word or not?
  - Whether it contains the sound /f/?
  - Whether it belongs to the category of mammals?
- What properties of words influence reaction times?
  - Length, frequency, complexity etc.
- This is easily measured by computers (not always that easily)

**Priming**
- In almost all areas of psycholinguistics
- If two things (representations) are related, they engage the same level of processing, and then processing one will have an effect on processing the other.
• This effect can be facilitatory or inhibitory. The relationship can be semantic, morphological, phonological, associative etc.

  *bird* → *thrush* versus *mammal* → *thrush*
  *bread* → *butter* versus *shoe* → *butter*
  *punishment* → *punish* versus *bulletin* → *bullet*

**Observation, surveys and questionnaires:**
• Mainly in child language research
• CHILDES database
• Diaries of (language researcher) Mums
• Important in research on language pathologies also
• Speech errors

**General methodological problems:**
• Most experiments tests monolingual typical university students (WASP)
• Most experiments involve reading, while for most people, speech is the central language activity
• How sensitive is performance to a specific experimental technique?

**A well-controlled experiment:**
1. Enough items
2. Matching relevant factors
3. Enough subjects
4. Statistical testing
5. Outcome task-specific?
6. Outcome language-mode specific?
7. Outcome language-specific?
2. BRAIN AND LANGUAGE:

Neurolinguistics:
- Where are specific linguistic functions localized in the brain? What are the anatomical structures and physiological processes that play a crucial role?
- How does the brain process and produce language?
- Localist versus distributed views

How can we examine language in the brain?
1. Utilization of the anatomical arrangement of different sensorimotor systems to study left vs. right hemisphere involvement in language tasks
2. Detailed analyses of patients with circumscribed brain damage and subsequent language disorders
3. Measurement of brain activity during performance of a language task
4. Producing a temporary, local disturbance in brain function while performing a language task

Lateralization of language:
- In principle, language is left-lateralized in the brain
- Women more likely to have bilateral language representation than men, but the results are somewhat mixed
- Handedness also plays a role:

<table>
<thead>
<tr>
<th></th>
<th>Left-sided language</th>
<th>Right-sided language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right-handers</td>
<td>~95%</td>
<td>~5%</td>
</tr>
<tr>
<td>Left-handers</td>
<td>~70%</td>
<td>~30%</td>
</tr>
</tbody>
</table>

Overall, for ~93% of the population, left is dominant for language.
- Possible determinants: genes, certain hormone levels pre- or postnatally

2.1. Methods that utilize the anatomy of sensorimotor systems in the study of hemispheric lateralization:
- Tachistoscope in the study of the visual language system.
- Dichotic listening in the study of the auditory language system.
- Based on the crossing pathways from the ears to the primary auditory cortices and from the eyes to the primary visual cortices: The main input goes to the contralateral hemisphere.
- By comparing the processing efficiency of stimuli “presented to each hemisphere”, it is possible to make inferences about the dominance of functions in them (e.g. language).
**Tachistoscope (the visual half-field technique):**

- Based on the crossing of the visual pathways (see figure): the main destination of the left visual field of the eyes is the right visual cortex etc.

*Split brain*

E.g. a word is presented in the left visual field or in the right visual field of the participant: which one is processed faster?
Probably the one presented in the right visual field (-> left hemisphere).
- But not the same for emotional words!
- Normally, the corpus callosum can communicate the information also to the other Hemisphere.
- In Split Brain patients (whose corpus callosum has been cut), the communication is not possible.
Visual field and hemispheres:

Kimura (1973):
Right visual field: better identification

<table>
<thead>
<tr>
<th>MODALITY</th>
<th>LEFT/RIGHT HEMISPHERE TEST SCORE RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory</td>
<td></td>
</tr>
<tr>
<td>words</td>
<td>1.88:1</td>
</tr>
<tr>
<td>nonsense syllables</td>
<td>1.73:1</td>
</tr>
<tr>
<td>backward speech</td>
<td>1.66:1</td>
</tr>
<tr>
<td>melodic pattern</td>
<td>1:1.19</td>
</tr>
<tr>
<td>human nonspeech sounds</td>
<td>1:1.08</td>
</tr>
<tr>
<td>Visual</td>
<td></td>
</tr>
<tr>
<td>letters</td>
<td>1.23:1</td>
</tr>
<tr>
<td>words</td>
<td>1.47:1</td>
</tr>
<tr>
<td>two-dimensional point location</td>
<td>1:1.18</td>
</tr>
<tr>
<td>dot and form enumeration</td>
<td>1:1.20</td>
</tr>
<tr>
<td>matching of slanted lines</td>
<td>1:1.05</td>
</tr>
<tr>
<td>stereoscopic depth perception</td>
<td>1:1.28</td>
</tr>
</tbody>
</table>
Split brain: (R. Sperry, M. Gazzaniga)

Name the object!

Patient: Spoon!

Name the object!

patient: (doesn’t say anything)
researcher: have you seen anything?
patient: No.
Rationalization (Choose the object that goes with the picture):

Chimeras:

Dichotic listening:

**Auditory Input:**

- word presented to the LH will be perceived (Dog).

“Right Ear Advantage”

The word presented to the right ear is perceived better than the one to the left: auditory language left-dominant.
Lateralization of Language:
**LEFT hemisphere:**
- Analytic processing
- Better ability to process temporal aspects of language
- Sequences of units
These aspects are very important in language processing, e.g., understanding of speech. Some disorders of language may be due to problems with timing ability.

Yet, some aspects of language are typical for the **RIGHT hemisphere:**
- Holistic processing
- Prosody: melody, tone of voice, stress
- Emotional aspects of speech
- Understanding of humour, metaphor, moral
- Capable of understanding simple speech

**Neural components of different levels of processing:**
Orthographic
Phonological
Morphological
Syntactic
Semantic

History of Neurolinguistics:
- References to the 'loss of speech' in the Smith Surgical papyrus (i.e. 1500-3000)
- Hippocrates: paralysis and loss of speech go together (i.e. ~450 I.E.)
- Gall (1758-1828)
  - Theory of phrenology
  - Beginnings of brain mapping
  - Different areas in the brain have different functions
  - The shape of the head and skull above these reflects how well-developed a certain function is.
Speech and motor systems:
- Speech involves cca. 100 muscles
- Motor behavior in primates is controlled by at least 3 distinct motor systems
  - separate movements of digits
  - Independent movements of arms and hands
  - Posture and bilateral body and limb movements
- It is probably the first that is responsible for speech

Lateral view of the human brain:
Hemispheres:
- Two hemispheres divided by a long fissure
- Communication between hemispheres through corpus callosum
- A wide bundle of neural fibres

The four lobes:
Functions generally attributed to different lobes:
- **Frontal**: motor coordination, executive functions, language, memory
- **Temporal**: auditory processing, language, memory
- **Parietal**: processing somatosensory info, spatial cognition
- **Occipital**: visual processing
Planum temporale:
- Asymmetric in humans (less so in women and left handed people)
- Important in speech processing
- Or, more generally in processing complex sounds (bigger in people with absolute pitch [absolutes Gehör])
Important subcortical areas in language: a basal ganglia and cerebellum

B = Broca’s area
W = Wernicke’s area
A = Auditory cortex
M = Motor cortex
Ss = Somatosensory cortex
Sm = Supramarginal gyrus
AG = Gyrus angularis
TP = Temporal pole
It = Inferotemporal cortex
Pf = Prefrontal cortex
Aphasia:
- The ability of language use or comprehension is impaired
- In most cases the damage is in the left hemisphere
- Most frequent cause: stroke (25 - 40% of stroke survivors show aphasia). Head injury, brain tumor, or other neurological reason
- Aphasia is most severe right after the injury, and gets better with time – though in many cases, speech and language problems are persistent
- Prevalence: little less than 0,5%

2.2. Study of brain damaged patients (Paul Broca 1824 – 1880)

Damage in Broca’s area
- Problems in production: articulation, poor use of grammatical features, word finding difficulties Understanding of speech fairly normal
- Plasticity of the young brain

Paul Broca and Leborgne:
- 1861: Leborgne (TAN)
- Sudden loss of speech at 21 – (only the syllable “tan”)
- Right hemiparesis
- Fairly intact comprehension and cognition
- Autopsy: a hole in the left inferior frontal lobe (BA 44/45; syphilis)
- Broca: the location of “the special ability of articulated speech”
- More cases like that in Broca’s descriptions
- Nonfluent aphasia: Broca’s aphasia—most cases are not as severe as Leborgne
Leborgne’s brain:

Top right: Close up of the brain, showing the area Broca identified.
Bottom left: Dr. M. Berthoz about to conduct MRI of the brain.

Carl Wernicke 1848 – 1904:

Damage in Wernicke’s area
- Prosody and pronunciation intact, speech is fluent but “empty”, but a lot of different word distortions and difficulties finding the right word
- Severe comprehension deficits
Classification of aphasias:

<table>
<thead>
<tr>
<th></th>
<th>Fluency</th>
<th>Comprehension</th>
<th>Repetition</th>
<th>Naming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broca’s</td>
<td>--</td>
<td>+</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Wernicke’s</td>
<td>+</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Conduction aphasia</td>
<td>+</td>
<td>+</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Anomia</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>--</td>
</tr>
</tbody>
</table>

“Could you tell me what happened?”

→ BROCA’

“Alright... Uh... stroke and uh... I ... Huh tawanna guy....h...hot tub and... And the ... two
days when un... hos... uh...huh hos-pital and uh... amet... am...ambulance.”

→ WERNICKE

“It just suddenly had a feffort and all the feffort had gone with it. It even stepped my horn.
They took them from earth you know. They make my favorite nine to severed and now
I’m a been habed by the .... Uh.... stam of fortment of my annulment which is now
forever.”

The classical view of language in the brain:

- Two language centres, for production and comprehension respectively:
  Broca’s and Wernicke’s area.
- The arcuate fasciculus: a bundle of nerve fibers connecting Wernicke’s area to Broca’s, is
  essential for normal language function. Damage to it causes conduction aphasia:
  speech “fluent”, auditory comprehension relatively good, but repetition of heard words is
  impaired.
The Wernicke-Geschwind model:
• The most well-known classical view of language
• Language processes basically progress from the posterior part of the left hemisphere towards the frontal parts
• High-level planning and semantics more in the posterior parts
• Low-level sound retrieval and articulation more in the frontal lobe

Classical localization theory
• Broca’s aphasia: frontal
• Wernicke’s aphasia: posterior
• Conduction aphasia: fasciculus arcuatus

Classical theory
Broca’s and Wernicke’s area in modern textbooks

Changes in 1960’s:
• The rise of generative grammar: language is autonomous
• „Mental organs“ of language
• Aphasia: impairment of the language organ
• Left frontal regions previously associated with motor language functions are now specifically associated with grammar
• Posterior areas are now associated with semantic function

Cognitive Neuropsychology Double Dissociation:
Double dissociation:
1. Different functions
2. Different brain areas
E.g. regular versus irregular suffixation (grammar vs. lexicon) or reading regular English nonwords (SPUKE) vs irregular existing words (STEAK)

Other language areas:
- **Dorsal thalamus**: dysarthria and aphasia
- **basal ganglia**: Parkinson’s and Huntington’s
- **Cerebellum**: dysarthria and ataxia

Rethinking the classical theory
1) Do patients show the symptoms predicted by the classical theory?
2) Do lesion-symptom pairs match predictions?
3) Is aphasia specific to the linguistic domain, as predicted by classical theories?
- Problems with testing neuropsychological patients: impairment of one area can affect functioning in other areas.
- Imaging techniques find metabolic abnormalities in otherwise healthy tissue at a distance from the lesion site.

The modern view of language in the brain
- Broca’s area and Wernicke’s area not homogenous pieces of tissue: can be subdivided into functionally distinct parts.
- Damage in Broca’s or Wernicke’s area does not correspond strictly to the aphasia syndromes the classical model predicts.
- Broca’s area activated also in *input* processing of language (not only production), The left inferior frontal cortex (incl. Broca’s area) may play a language executive role in coordinating phonological, syntactic and semantic processes in posterior areas and inhibiting irrelevant information.
- Wernicke’s area is important in auditory processing but is not the only location where language comprehension occurs (this involves several regions).
- There are important language-related areas outside the classical ones: e.g., middle and inferior sections of the temporal lobe, inferior parietal cortex, basal ganglia and cerebellum.

Broca’s aphasia and TAN
- Cases like that are very rare
- They are diagnosed as Broca’s
- But impairment in Broca’s region is not systematically connected to this deficit type
- Dronkers et al: this type of problem is more often related to a lesion disconnecting the arcuatus fasciculus

Is there no point in trying to localize linguistic functions?
- Maybe there is.
- Consistent lesion sites are difficult to find in Broca’s and Wernicke’s because these are syndromes not symptoms
- Language functions can be disturbed in many ways and it is not likely that impairments are related to one or some brain areas
- The more fine-grained a behavioral measure is, the more likely successful localization is.
Nonlinguistic impairments in aphasia

- “slow thinking” (Jackson, 1878), “asymbolia” (Finkelnburg, 1870).
- Apraxia and other impairments in production and comprehension of actions
- Nonlinguistic impairments have been demonstrated:
  - deficit in picture-object mapping (De Renzi et al., 1968b), in gesture-object mapping (De Renzi et al. 1968a), sound-picture mapping (Schnider et al., 1994; Spinnler & Vignolo, 1966; Varney, 1980)
- More recently: processing nonlinguistic environmental sounds, and pantomime (Saygin et al, 2003a,b)
- Do these impairments systematically cooccur with aphasia?
- What are their behavioral correlates?

2.3. Measurement of brain activation during language tasks:

1.) Methods to measure hemodynamic variation:

- PET: Positron Emission Tomography
  - Radioactive water is inserted into blood, and the concentration can be detected outside of the brain by the PET scanner.
  - Blood flow (and thus also the degree of radioactive water) is increased in brain areas that are active during an experimental task, which allows the localization of this function to certain brain areas.
  - Good spatial resolution (good for localizing functions to certain brain areas), poor temporal resolution (not good for studying the timing of processes)

- fMRI: functional Magnetic Resonance Imaging
  - Based on blood oxygen level and its magnetic properties: active areas consume more oxygen, and this is (indirectly) measured during a cognitive task.
  - High spatial resolution (= good for localizing), poor temporal resolution (= poor for investigating the timing of processes)

Example:
Task: translating silently visually presented sentences from Finnish into Norwegian
Control task:
reading silently visually presented Finnish sentences Subtraction -> Brain areas activated specifically for translation (across-language retrieval of words and sentence structure)
2) Methods to measure electromagnetic activity:

- EEG: Electroencephalography
  - Based on the recordings of electrical brain activity measured at the surface of the scalp.
  - Good for studying the *timing* of cognitive processes:
    - time resolution excellent,
    - spatial resolution poor

**Example:**

Presenting *semantically* anomalous Sentences Elicits a negative peak ~400 ms after presentation of the unsuitable word (“N400”).

Presenting *syntactically* anomalous Sentences elicits a positive peak ~600 ms after presentation of the unsuitable word (“P600”).

The brain makes a distinction between syntax and semantics.
N400: semantic integration
P600: syntactic integration
- **MEG**: Magnetoencephalography
  - Measures magnetic fields produced by electrical activity in the brain.
  - Fairly good spatial resolution, excellent temporal resolution (can be used both to studying the time-course and localization of language processes)
  - Possible to work out the sequence in which different brain areas contribute to processing.

2.4. **Methods to locally disturb brain function:**

- Sodium amytal procedure (Wada test)
- Cortical stimulation
- TMS: transcranial magnetic stimulation

**WADA test:**

- Sodium amytal...
- **Juhn Atsushi Wada**
2.5. **Neural components of different levels of processing:**

Neural correlates of **orthographic** processing:

![Diagram showing neural components of different levels of processing](image)

**Functional Neuroimaging Studies:**
- Left ventral occipito-temporal areas, complex visual processing

An early PET study
Neural correlates of **phonological** processing:

**Lesion Studies:**
Conduction aphasia
- repetition disproportionately severely impaired
- Main problems in the proper choice and sequencing of phonemes in speech output, leading to numerous **phonological errors**

▶ **A Hypothetical Explanation of Conduction Aphasia**

**Functional Neuroimaging Studies:**

---

**Fig. 1.** A simple model of the cortical network supporting speech perception and related language functions. The dashed line indicates the possibility of additional, non-parietal auditory-motor interface networks (see text).
Neural correlates of **morphological** processing:

**Lesion Studies:**
- Inflectional and derivational morphology, and compounding can be selectively impaired in aphasia
- Patients with predominantly *temporal* lobe lesions have more trouble producing irregular (e.g., drive-drove) than regular (talk-talked) verb forms.
- Patients with predominantly *frontal/basal ganglia* lesions have more trouble producing regular than irregular verb forms.

Neural correlates of **syntactic** processing:

**Lesion Studies:**
- at word level, patients with left frontal damage can have more trouble with verbs than with nouns, while patients with left temporal damage can have more trouble with nouns than with verbs

- at sentence level, agrammatism / morphosyntactic difficulties in *production* has been associated to damage in left frontal areas (e.g., Broca’s)

- at sentence level, morphosyntactic difficulties in *comprehension* have been associated with a variety of left perisylvian lesions
Functional Neuroimaging Studies:

- Broca’s area activation for many tasks involving syntactic processing: e.g., processing of complex vs. simple sentences (see figure), and syntactic violations
- However, many other areas are relevant as well!
- No unique area for syntax, different parts of the network are recruited for different aspects of processing (The same seems to be true for most other subcomponents of language)

![Brain imaging](image)

Processing syntactically complex vs. simple sentences

Neural correlates of semantic processing:

Lesion Studies:
- Wernicke patients can have problems with semantic tasks
- Some dementia patients exhibit a slowly progressive condition coined as “semantic dementia”. It is associated with left temporal lobe damage.
- There can be category-specific naming disorders -> distinct areas for different categories of concepts, e.g., for tools and animals.

![Brain diagram](image)
Functional Neuroimaging Studies:

Brain regions activated during various lexical-semantic processing tasks include, but are not limited to the following:

- left inferior frontal lobe
- left inferior parietal – posterior superior temporal
- left fusiform/inferior temporo-occipital regions
3. LANGUAGE AND THOUGHT

3.1. Relationship of cognition and language

- Categories of cognition are shaped by language
  - Sapir and Whorf’s linguistic relativity
- Cognitive categories develop independently of language both in evolution and in ontogeny, language only builds upon these
  - Piaget: cognitive development leads language development
- Language and cognition are independent – Chomsky
- Cognition follows its own path, but language modulates its categories

3.2. Linguistic relativity

- ’50-s: big differences, mainly relativity
- Chomskyan revolution: universalism and modularity, no determinism
- ’80s: typological differences, parameters, processing differences
- Today: constrained relativity and contextualized universality

Whorf and linguistic relativity

- language determines thought
- This is manifest in differences in expressions of different languages lexical and grammatical relativity
- Three parts to study: language, thought and ‘determines’
- **linguistic determinism**: a language shapes psychological mechanisms
- Benjamin Lee Whorf
  - Language shapes the mind, world view, structure of science
  - Differences in lexical (vocabulary) and grammatical organization result in different conceptual schemes
- Cultural anthropologist in the Boas and Sapir tradition
  - Emphasized the variety and differences of cultures, not the common features
- Strong view: all higher forms of thought build on language
- Weak view: the structure of the language one generally uses, influences the way they understand their environment and act upon it

Linguistic relativity

- Follows from linguistic determinism
- Linguistic relativity: distinctions encoded in one language are unique to that language alone, and that “there is no limit to the structural diversity of languages” – translation!
- lexical and grammatical relativity

Lexical influences

- Lexical level:
  - the matter of what words are found in a given language, and what they refer to
    - different languages carve up the reality in different ways through more or less specialized vocabularies
    - languages differ with respect to how they divide up the world into nouns and verbs
      - lightning: a N in English, but a V in Hopi
        - duration an important feature
• Tzotzil Mayan: eat-mushy; eat-a-slender-shape-food, eat-meat
  — the properties of objects are incorporated into the verbs

**Grammatical influences on thinking**

• Number category
  — whether inanimate nouns can be pluralized or not
  — in English can pluralize any noun as long as the referent is discrete, i.e., mass nouns such as *paper, flour* cannot be pluralized - count nouns such as *pen, girl*
  — in Yucatec, only animate nouns can be pluralized
  — Lucy (1992): English speakers specify the number of objects in descriptions of line drawings more frequently than Yucatec speakers

• Tense markers
  — determine location of events in time
    past -------- now -------- future
  
  - he is running
  - he ran *War!* in Hopi
  - he will run

• How does a temporal language compare to a “timeless language”?
  Tense:
  • Hopi distinguished between
    - report of an ongoing event (visual field non-blank)
    - report of an event from memory or expected event (visual field blank)

• Whorf: these distinctions affect the way cultures are organized and how science is done
• Is language a conceptual prison?

**Examples**

1. snow
2. colours
3. gender
4. Spatial language

**1. Snow:**
Eskimos have many different words for ‘snow’ → evidence that they see snow differently *(urban legend!)*

→ Boas, (1911): 4
  • aput („snow on the ground“)
  • gana („falling snow“)
  • piqisirpoq („drifting snow“)
  • qimuqsuq („a snowdrift“)
→ Sapir & Whorf, 1940: 7
→ 1978: 50
→ 1984 *(New York Times): 100*

The truth about snow
• There are several eskimo languages + eskimo laguages differ in the number of expressions they have for snow
• English has several expressions for snow
• ~same number of snow-expressions in English and Eskimo
• Definition of „word” is problematic - inuit: are words derived from the same stem different or not?

BUT
Even if it was true that one language had more, is it evidence that they see snow differently?
painters: paints
ornithologists: birds

2. Colors

Basic colour terms (Berlin & Kay, 1969)
• Properties:
  – 1 morpheme
  – Not restricted to one class of items (e.g. blond)
  – Do not belong to the scope of another color terms (e.g. turquoise)
  – Frequently and generally used
• Basic color terms are chosen from 11 colors by all languages:
  black, white, red, yellow, green, blue, brown, pink, purple, orange, grey.
• languages differ in how many basic color terms they have
  (Hungarian for ‘pink’ rózsaszín is not a basic color term)

Languages: different set of basic color terms

<table>
<thead>
<tr>
<th>Languages</th>
<th>English</th>
<th>Purple blue green yellow red orange</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shona</td>
<td>Cipswuka citema cicena cipswuka</td>
</tr>
<tr>
<td></td>
<td>Bassa</td>
<td>Hui ziza</td>
</tr>
<tr>
<td></td>
<td>Dani</td>
<td>Mili mola</td>
</tr>
</tbody>
</table>

Different color perception?

Hierarchy of basic color terms

black, white

   red

yellow, green, blue

brown

purple, pink, orange, grey
Do speakers of different languages see colors differently?

- Color categories are not arbitrary!
- Same everywhere:
  - light
  - Operation of the human eye
- 3 kinds of cones in color perception → these determine what we see
- Experiments (pl. Heider, 1972 – the dani): recalling and discrimination is good for colors—focal colors

What is the origin of such great cognitive variety?

- Does it originate in linguistic differences?
  - Does language have a „whorfian” effect on cognition?
  - Language ▶ thought?
- Or in environmental factors?

Cognitive prerequisites view

- Piaget: conceptual development precedes linguistic (i.e. semantic) development
- Children first have to form concepts before they can form the linguistic terms encoding them
  - Object constancy < object names
  - Causality < causal connectives (e.g. because)
  - Spatial categories < spatial expressions

Evidence for cognitive prerequisites (Bowerman, 1996)

- Children’s spatial semantic concepts are rooted in nonlinguistic development
  - They say things like sit chair, ball table: they communicate about location of objects before they use spatial morphemes
  - Order of acquisition within language and between languages
    in, on < under < next to, between, in front of (egocentric) < behind, in front of (taking the orientation of target object into consideration)
  - They distinguish front and back of objects well before using in front of, behind, or front and back
  - They play with objects in a way that reflects understanding of the concepts of containment and support well before the appearance of in and on

Evidence against cognitive prerequisites (Bowerman, 1996)

- Korean (vs. English and Hungarian): no linguistic distinction, between placing an object in a container or on a surface (in vs. on, -ban vs. -on)
- Korean language distinguishes between tight fit (ring on a finger, picture on the wall) and loose fit (fruit in a bowl, object leaning against a wall)
  - This distinction holds for both containment (in) and support (on)
Development II.
(Spelke & Hespos, 2002)

<table>
<thead>
<tr>
<th></th>
<th>Container</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>the apple is in the bowl</td>
<td>the apple is on the table</td>
</tr>
<tr>
<td>Hungarian</td>
<td>az alma a tálban van</td>
<td>az alma az asztalon van</td>
</tr>
</tbody>
</table>

Korean:
*tight fit* (ring on a finger, picture on the wall)
*loose fit* (fruit in a bowl, object leaning against a wall)

This distinction holds for both containment (*in*) and support (*on*)
e.g. both letter in an envelope + magnet on fridge = both *kitta*

Are these differences reflected in spatial representations of English/Korean speakers?

Experimental results (McDonough et al. 2000)
– English/Korean babies differentiate all potential spatial distinctions
– As a result of acquiring a language certain spatial distinctions
  (those strengthened by language) become salient in representation
– (phoneme sensitivity in babies)
– language ► thought

Form Perception
• Carroll and Casagrande: Navajo vs. English
  – Navajo verbs change form according to the shape of the object it takes (shape classifiers)
  – flexible vs. rigid; flat vs. round
    – give blue rope and yellow stick and ask which of the two a blue stick can go with
• Navajo choose shape: yellow stick
  – English choose color: blue rope
  – conducted the test with upper class Bostonians
• responded like Navajo children
  – there is other kinds of determinism than just linguistic determinism

Hypotheticals
• Bloom: absence of conditional constructions in Mandarin Chinese
• can Chinese speakers think about counter-factual situations?
  – if he had arrived earlier, he would have seen us
  – less than English speakers
• Au: found contrary evidence to Bloom
  – direct empirical evidence for the Whorfian hypothesis has been disappointing

Interactionism: Semantics and Cognition
• Semantic development may facilitate conceptual development
  – language-specific knowledge may restructure our nonlinguistic knowledge
• Differences across languages might influence which conceptual distinctions children find easy or difficult to encode
Cross-linguistic work
• Examine development of concepts in languages that differ morphologically and/or grammatically in interesting ways
• Can sort out nonlinguistic development from what is language-specific

3.3. Language Evolution
• Continuity or discontinuity with other animal species
• Change in one single ability or more changes?
  – Concept formation and categorization
  – Articulation and speech perception
  – Syntax-motor organization—sequential and hierarchical
  – Social abilities
• Adaptation vs exaptation?
• Specific or more general abilities?
• Brain development
• Selection pressures: communication or representation
  – Group cohesion, courting, teaching, deception

Animal vs Human communication

<table>
<thead>
<tr>
<th>Feature</th>
<th>Animal communication</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation</td>
<td>similarity</td>
<td>Arbitrary</td>
</tr>
<tr>
<td>Inventory</td>
<td>Limited (50)</td>
<td>infinite</td>
</tr>
<tr>
<td>system</td>
<td>Closed</td>
<td>open</td>
</tr>
<tr>
<td>Flexibility</td>
<td>fixed</td>
<td>Variable</td>
</tr>
<tr>
<td>Situation</td>
<td>Stimulus-dependent</td>
<td>Independent</td>
</tr>
<tr>
<td>Control</td>
<td>automatic</td>
<td>intentional</td>
</tr>
</tbody>
</table>

Cultural transmission, double patterning
Animal communication

- Darwin: chimpanzee faces, disappointed

- Ritualization in wolves

- Dominant follows submissive, posture!

Animal vocalizations and human speech

- Vervet calls
  - Eagle, leopard, snake
  - Comprehension flexible, production not
  - Context-dependent
  - Not very sophisticated
  - Involuntary and inflexible
  
Communication signals (peacock’s tail as well)

- Primate gestures instead of vocalizations are the origin of human communication (Tomasello)
  - More flexible
  - More complex than in other species
  - Individual developmental pathways
  - Can be combined and used flexibly

Multiple patterning in human language

<table>
<thead>
<tr>
<th>Level</th>
<th>Function</th>
<th>Organization</th>
<th>Set size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonology</td>
<td>Distinction</td>
<td>Features-sounds</td>
<td>20-50</td>
</tr>
<tr>
<td>Morphology, words</td>
<td>Meaning</td>
<td>Sounds-words</td>
<td>Many thousands</td>
</tr>
<tr>
<td>Sentences</td>
<td>Declaration, action</td>
<td>Words-sentences</td>
<td>Infinite</td>
</tr>
<tr>
<td>Texts</td>
<td>Human relationships</td>
<td>Sentences-texts</td>
<td>Infinite</td>
</tr>
</tbody>
</table>
Recursion
- Mary dreamt that John asked why she told his friend that Frank scolded him for not closing the door when leaving to school to tell Kate that...
- The leg of the friend of the neighbor of the painter of the picture in my neighbor’s friend’s house broke

The road to language

Hauser, Chomsky and Fitch:
(FLB) Faculty of language—broad sense
- Continuous development
- Control of articulation
- Brain development
- Concepts
- Communicational efficiency
- Adaptive advantage

(FLN) Faculty of language—broad sense
- Human-specific
- Recursion
- Probably no adaptive advantage
**Biological changes on the road to language**
- Descent of larynx: longer cavity, more possibility for articulation
- More fine-grained sound discrimination
- More fine-grained motor organization
- Appearance of gestures
- Encephalization: bigger brain relative to body; growth of cortex, cerebellum
- Later speech centers: fine movements, gestures, sound
- Asymmetries in brain

**Changes in articulation: smaller chewing, longer sound channel (Philip Lieberman)**

![Diagram showing biological changes](image)
Evolution of the specific language ability
• Sudden change: Chomsky
  - Syntax sudden feature
  - No real advantage
  - Exaptation
  - Openness is the key
  - Species-specific

• Continuity Pinker
  - Syntax: gradual feature
  - Increases chance of successful courting
  - Adaptation, e.g. cooperation
  - Openness is the key
  - Species-specific

Jackendoff’s view

Box 1. Steps in the evolution of language
Independent steps appear side by side; dependencies among steps are indicated vertically.

Protolanguage \(\Rightarrow\) language (Bickerton)
• Early humans: imitation, tradition, action categories. Protolanguage: agent, goal, instrument, recipient, location categories
• Structural relations become formal (independent of content, syntax)
• Grammatical operations become productive (embedding, reuse of structures)
• Appearance of grammatical morphemes (suffixes, determiners, etc.)
• Word order used for functional roles
• Appearance of empty grammatical categories (pronouns, \textit{It is raining})
**Primates**
- Rich categorization
- No system
- Slow learning of symbols
- Restricted syntactic combination
- No tradition
- Few imitation

**Changes in architecture: primate representation (Merlin Donald)**

<table>
<thead>
<tr>
<th>Culture</th>
<th>Age</th>
<th>Knowledge type</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Episodic</td>
<td>primates, 5 m</td>
<td>events</td>
<td>No</td>
</tr>
<tr>
<td>Mimetic</td>
<td>erectus, 1.5 m</td>
<td>Bodily</td>
<td>Acting out</td>
</tr>
<tr>
<td>Mythical</td>
<td>sapiens, 50 e</td>
<td>Linguistic</td>
<td>Narratives</td>
</tr>
<tr>
<td>Theoretical</td>
<td>modern, 10 e</td>
<td>External memory</td>
<td>Writing-reading</td>
</tr>
</tbody>
</table>

**Cognitive changes and language**

<table>
<thead>
<tr>
<th>Cognitive function</th>
<th>Role in language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working memory</td>
<td>Word learning, sentence processing</td>
</tr>
<tr>
<td>Divided attention</td>
<td>Word learning, contextual interpretation</td>
</tr>
<tr>
<td>Brain plasticity</td>
<td>Lifelong learning</td>
</tr>
<tr>
<td>Increase in LTM capacity</td>
<td>Lexical storage</td>
</tr>
<tr>
<td>Extension of semantic brain regions</td>
<td>Richer contextual meaning</td>
</tr>
</tbody>
</table>

**Mithen**

- Special systems, than communication between systems
- Language and other cognitive systems
- Shift between h. erectus (2m – 200 thousand) and H. sapiens

---

Psycholinguistik – Folien, WS11  Seite 48
What was language initially good for?

- Hunting, organization of work
- Impressing mates: ‘peacock tail effect’, courting (Miller)
- Teaching and rapid classification (Hernánd)
- Dominance, deception: mind-reading
- Maintaining social relationships: gossip

Dunbar’s social theory

- Group size gets bigger with brain size
- Humans: 120 members
- With group size, grooming time increases
- With early humans, this would have taken up 40% of time
- First function of language is maintaining relationships, ‘grooming communities’
- Dialects and language variants create manageable group sizes within the large group

Grooming time and group size
Modern human relationships

• Social network max cca. 100
• Subgroups:
  - intimate: 7–10
  - medium: 20–40
  - broader: 100–400
• Emotional distance and social density

![Graph showing mean and standard error of months since last contact for different levels of emotional closeness.]

Figure 6: Mean (and standard error) number of months since last contact between individuals based on their degree of emotional closeness.

Mindreading, Michael Tomasello

Joint attention
Tomasello’ theory of the evolution of culture and sociality

<table>
<thead>
<tr>
<th>Activity</th>
<th>Basic social</th>
<th>Cultural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>Signals</td>
<td>Intersubjective symbols</td>
</tr>
<tr>
<td>Other’s gaze direction</td>
<td>Following gaze direction</td>
<td>Joint attention</td>
</tr>
<tr>
<td>Social learning</td>
<td>Following, ritualization</td>
<td>Reciprocal intentional actions</td>
</tr>
<tr>
<td>Cooperation</td>
<td>Coordination</td>
<td>Role differentiation</td>
</tr>
<tr>
<td>Teaching</td>
<td>Facilitation</td>
<td>Instruction</td>
</tr>
<tr>
<td>Object manipulation</td>
<td>Tools</td>
<td>Intentional tool use</td>
</tr>
</tbody>
</table>

Basic mindreading to language

Following gaze direction → Persuasion

\[
\text{A P belief} \quad \text{B P Deception}
\]

Joint attention

\[
\text{A P belief} \quad \text{B Q}
\]

Word learning

Miller: sexual selection
- Mating is central motivation
- Features of attractive mates spread
- Features indicating fitness are attractive
- Speech is a fitness indicator, because it is complicated and superfluous
- Peacock-tail effect
Reproductive success above all (runaway selection)

Overall picture

- Primates: motor organization, vision, individual learning, categories
  - Increase in brain size, gradual gesture use, first social learning 2 m
  - Inventory of sounds gets richer. Tools, general categories, Protolanguage. 1 m – 100 e
  - Innovative tool use. FOXP2 mutation: formal grammar 100-50 thousand d years
4. LANGUAGE ACQUISITION

The task:
- Acquisition of phoneme inventory
- The segmentation of the sound wave into
  - Phonemes
  - Morphemes
  - Words
  - Sentences
- Assignment of meaning to units
- rule extraction
- categorization and labeling (bootstrapping)
- exceptions

4.1. Stages

**• Prelingual stage (0 – 12 months)**

The acquisition of the phonological system during the first year of life: perception and production

Crying Birth
Cooing 6 Weeks
Babbling 6 Months
Intonation 8 Months

**• One-word (holophrastic) stage (12 – 24 months)**

- Name people, objects, etc.
- An entire sentence is one word
- phonological, semantic errors
- The acquisition of lexical meaning
- Vocabulary spurt (1.5 years)

**• Two-word stage (24 – 36 months)**

- Two-word and later multi-word utterances.
- Use consistent set of word orders: N-V, A-N, V-N...
- With structure determined by semantic relationships
  - agent+action (baby sleep)
  - possessor+possession (Mommy book)
- Telegraphic stage (only content words)

Pivot-Open 18 Months
Word Inflections 2 Years
Questions & Negatives 2 ½ Years
Complex constructions 5 Years
Mature Speech 10 Years
4.2. Methods:

- Prelinguistic
  - Measurement of heart beat (prenatally)
  - High Amplitude Sucking Procedure (shortly after birth)
  - Preferential looking paradigm (from six months)
- Diaries
  - First words
- Standardized parental reports (CDI)
- Unconstrained speech recordings (corpus, CHILDES) – longitudinal studies
  - Phonological development
  - Vocabulary development
  - Pragmatic and discourse development
  - Comparisons of parental and child language
  - Grammatical development
- Experiments—cross sectional
  - Language comprehension
  - Specific grammatical phenomena
  - Grammaticality judgments

Methods: sucking rate:

A 1-month-old infant sucking on a special nipple while listening to a recording of different syllables. The child’s normal sucking rate is first established, and then sequences of sounds are presented. As the child hears the first sound, the rate of sucking increases. During subsequent repetitions of the same sound, the sucking rate shows a gradual decrease. A new sound is then played to the child. If no distinction is perceived, the sucking rate will continue to decrease; but if a change is perceived, it will show a sudden increase.
Methods: preferential looking:

Hey! She's kissing the key!
Elicited production with nonsense words:
Berko 1958 wug-test:

- Prototype of language acquisition experiments
- Evidence that language acquisition is structure building

Child-directed speech:
- Simplified Vocabulary
- Simplified Phonology
- Exaggerated pitch, intonation and duration
- Many Questions by Mothers
- Many Imperatives by Fathers
- Baby-Talk Words
  e.g. wawa, choo-choo, tummy, scrambled eggs, pasghetti

4.3. The acquisition of sounds:
- 0-2 months: crying
- 2-6 months: cooing (especially in the back of the mouth).
- 7-11 months: reduplicated babbling (repetitions of CV sequences).
- 11+ months: variegated babbling
  (more consonants and vowels, with sentence-like intonation).

Properties of easy sounds:
Front of the Mouth
Total Articulation
Muscles already Developed (in Nursing)
Easy Sounds: /m, p, b, t, d/
Hard Sounds: /ŋ, Θ, ð, s, r, l/ clusters
Easy sounds occur in more languages and are learned earlier by children.
Effects of prenatal exposure:
• Prenatal exposure: mother’s voice (filtered sound)
• Recognition of a prose passage before birth (DeCasper et al. 1994)
• Recognition of the native language (after birth (Mehler et al. 1986)
• Early speech sound discrimination
  • not limited to mother tongue
• Categorical perception (Eimas et al. 1971).
• Decline in discrimination of nonnative vowel contrasts between six and eight months (Polka and Werker)
• Decline in discrimination of nonnative consonant contrasts between ten and twelve months (Werker and Tees 1984).
• Other aspects of the phonological system of the native language (Jusczyk):
  phonetic inventory, predominant stress pattern, phonotactic regularities, word segmentation (stress influence).

4.4. One-word stage: lexical development:
• Onset around nine months (comprehension) or 12 months (production)
• Begin with simple lexical items for people/food/toys/animals/body functions
• 14,000 words around the age of six (approximately 10 words per day after the age of two).
• Between six and 17: approximately 3,000 words per year: 50,000 words at the age of 17.

Individual differences in vocabulary size:

<table>
<thead>
<tr>
<th>Age</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 months</td>
<td>10</td>
<td>110</td>
</tr>
<tr>
<td>18 months</td>
<td>13</td>
<td>230</td>
</tr>
<tr>
<td>20 months</td>
<td>25</td>
<td>380</td>
</tr>
<tr>
<td>22 months</td>
<td>40</td>
<td>470</td>
</tr>
<tr>
<td>24 months</td>
<td>50</td>
<td>510</td>
</tr>
<tr>
<td>26 months</td>
<td>110</td>
<td>600</td>
</tr>
<tr>
<td>28 months</td>
<td>180</td>
<td>590</td>
</tr>
<tr>
<td>30 months</td>
<td>340</td>
<td>650</td>
</tr>
</tbody>
</table>

Mean Length of Utterance (MLU)
Lexical development starts slow

![Graph showing the relationship between age and MLU in morphemes for 123 children.](image)

**Figure 10-1.** Relationship between age (± 1 month) and mean length of utterance (MLU) in morphemes for 123 children. [From J. F. Miller and R. S. Chapman, “The Relationship between Age and Mean Length of Utterance in Morphemes,” *Journal of Speech and Hearing Research* 24 (1981): 154–161.]
Gopnik and Meltzoff
- Strong causal connection between nonlinguistic categorization abilities and word learning (exhaustive sorting)
- Engine of vocabulary spurt is a more general learning mechanism: fast mapping (Carey and Bartlett 1978)
- The child learns initial information concerning word meaning after a few or even a single exposure

Lexical development
- Locke (1690)
- “… to make children understand what the names of simple ideas or substances stand for, people ordinarily
  1. show them the thing whereof they would have them have the idea
  2. repeat the name that it stands for “
- Challenges for Locke’s Theory: the child is NOT looking at the elephant while hearing the word “elephant”, but at adult’s face (not the cookie).
  Mapping error?
- Adult face = “elephant”
4.5. Vocabulary acquisition

vov-vov dog
for dogs, kittens, hens, zoo animals
mooi moon
for moon, cake <O> anything round
dany bell sound
for bell, clock, telephone, doorbell
quack duck sound
ducks, birds, insects, coins (because a coin had an eagle on it)
koko rooster crowing
rooster, merry-go-round, musical sounds, all sounds

Noun bias
• Nouns often refer to individual objects, persons, etc. Therefore, their meaning is relatively easy to figure out.
• Verbs, on the contrary, are relational words, i.e., you cannot imagine what a verb means without the verb's arguments. Therefore, their meaning is relatively difficult to figure out.
• It is not an input effect because:
  – Estimates say that verbs are more frequent than nouns.
  – Even in 'verb-friendly' languages (e.g., Korean, where verbs come at the end of utterances, and occur often in isolation) children start with nouns and not with verbs (?)

Acquisition of words
• Idiomorphs
• Nominals (words for objects, persons, animals) are more common than verbs (referential versus expressive children).
• Phonological errors (input effect?)
• Semantic errors (overextensions and underextensions)
Then:
• children come to use words in more adultlike ways
• words start to be used in wider range of contexts
• children use wider range of word types:
  – referential words (ball, doggie, chair)
  – proper names (Mummy, Spot)
  – actions (open, wash, tickle)
  – properties, states, qualities (more, gone, up, on, dirty)
  – social-pragmatic words (no, please)
  – few ‘frozen’ phrases (all gone, what’s that)

Lexical development
• First words are basic level categories
  – Dog or flower < poodle, tulip or animal, plant
  – Reflects adult frequency, functionally most useful level
• Differences from adult use
  – Overextensions: A word is used for something that has a similar shape, color, or function as the original referent (e.g., ball for moon, dog for bear, donkey, wolf, etc.).
  – Pure overextensions are those that are based on one aspect only, whereas mixed
overextensions are those that are based on more than one aspect.
– Overextensions based on appearance and shape are much more common than those based on function (e.g., car for a sled).
– **Underextensions**: A word is used for only a subset of what the word refers to (e.g., shoe for the child's own shoes but not for someone else's shoes).

**Knowing a word**

Estimated size of the adult lexicon: 50,000 -- 100,000. For each of these we know its
– meaning
– How its form and meaning can be changed
– How frequent it is and how often it occurs together with other words
– How we can combine it into phrases, clauses and sentences
– We know the pragmatic considerations of its use

**Representations of word meaning**

**Semantic feature theory (Clark, 1973, 1975):**

• each word has a list of semantic features:
  – e.g. DOG = +OBJECT +ANIMATE +FOURLEGGED+ FURRY +WHISKERS +WOOFS
  – a referent (object) must be characterized by all these features for the word to be applicable
  – children start with more general features (e.g. +OBJECT +ANIMATE) then extend to include more specific features (+WOOFS) later on
• **Good points:**
  – explains OVEREXTENSION errors
e.g. daddy = all adult males
  – explains why overextension tends to apply to perceptually similar shapes
• **Bad points:**
  – overextension not as frequent as Clark thought. Barrett (1996) - 7-33% of words
  – overextension occurs late in the developmental history of a word (Dromi, 1987)
  – underextensions more common early on (Golinkoff et al, 1994)
  – it’s proven impossible to define the relevant sets of semantic features
  – how does this work for verbs and other words (e.g. close)

**Prototype theory**

• Widely supported
  – Meaning of a referential word is initially acquired in the form of a prototypical referent for that word
  – e.g. meaning of word dog 1st applies only to a typical dog
  – child then generalises to other objects on basis that they share common features with the prototype then
• **Good points:**
  – explains overextension
  – e.g. clock -> bracelet AND sound of dripping water
  – explain underextension
• **Bad points:**
  – no one agrees as to what is prototypical
  – can’t explain acquisition of non-referential words
  – can’t explain why initial words occur in restricted range of contexts (Barrett, 1986)
The induction problem
• Quine, 1960, a visitor to a foreign land encounters a native who, looking towards a rabbit hopping through a field, says “Gavagai!”. To what does gavagai refer?
• It could mean “rabbit,” “hopping,” or “furry”. It could mean “Let's hunt that”, or “Death awaits you all with nasty, big, pointy teeth” (Gilliam & Jones, 1975), etc.
• The utterance “Gavagai”, by itself, is too underspecified to refer to anything particular in the world of an unconstrained learner.
• “For, consider ‘gavagai’. Who knows but what the objects to which this term applies are not rabbits after all, but mere stages, or brief temporal segments, of rabbits? In either event the stimulus situations that prompt assent to ‘Gavagai’ would be the same as for ‘Rabbit’

Solutions
Constraints on word learning?
– Domain-specific constraints on possible meanings e.g. Markman, 1987)
– General processes: associative learning, communicative intent (Tomasello, 1998) or knowledge of theory of mind (Bloom, 2001)
– A combination of the above (Golinkoff, Hirsh-Pasek et al., 2000)

Nativist solutions
• Children have innate knowledge that enables them to learn words
• Constraints theory (Markman, 1989, 1992, 1993)
  – built-in assumptions direct mapping of words onto meanings
• CONSTRAINTS:
  – WHOLE OBJECT CONSTRAINT
  – TAXONOMIC CONSTRAINT
  – CONTRAST CONSTRAINT
• similar constraints for actions(Clark, 1993)

Markman (1989)
• Children are biased to entertain certain hypotheses about word meanings over others
• These first guesses save them from logical ambiguity, and get them started out on the right track
• Mutual exclusivity
• two words must refer to two different things
  – Early vocabularies consist of basic level words that are mutually exclusive.
  – Children say things like: That is not a car it is a taxi.
  – Experimental evidence: Children select an unfamiliar object when they hear a novel word
• Whole object
  - A new word refers to the whole of a novel object, not its parts
• Taxonomic
  - a new word refers to a class of novel objects, not a specific individual
Taxonomic Categorization Bias:
- **Taxonomic relations**: tokens of the same type (e.g., dog and cat).
- **Thematic relations**: objects that have a spatial or a causal relation, or that often occur together in the same situation (e.g., spider and web).
- Words in languages indicate taxonomic rather than thematic relations (*why*?).
- Do children search for thematic relations or taxonomic relations (Markman and Hutchinson)?

**Clark (1983): pragmatic principles on lexical acquisition**
- **Conventionality**: every word has a meaning, shared by all speakers of the language
- **Contrast**: the meaning of a word contrasts with the meaning of all other words in the lexicon

**Lexical Contrast**
- Original study: Kagan (1981), as part of a large scale study on development in the second year
- Children were shown 3 objects, 2 familiar (a doll and a dog) and one unfamiliar
- ‘Give me the zooob!’
- By 22 month of age, children tended to chose the novel object →
  They believe that a novel word does not refer to objects that already have a name
- Methodological problem: children repeatedly tested over several months

**Markman and Wachtel (1988)**
- Preschool children shown a familiar (banana) and a novel object (whisk)
- ‘Show me the fendle!’
- They tend to interpret the new word as naming the whisk
- If only the banana is present, children tend to interpret the new name as an object part.
  → Evidence for the application of the mutual exclusivity principle

**Mutual exclusivity**
- ME applies only relative to one particular language: Au & Glusman (1990):
  if a Spanish-English bilingual knows the name of something in English, readily accepts the second name if it is given in Spanish
- Only a default assumption, but not immutable (Markman 1992)—languages frequently violate mutual exclusivity (*dog, puppy, pet, animal*)
- Specific to lexical acquisition (either innate or acquired in the course of word learning)
- Special case of a general principle of learning, one guiding children to prefer one-to-one mappings as a general tendency to exaggerate regularities (Markman, 1992)
- ME is the product of children’s theory of mind (Bloom, Clark)
- Infants reason as follows:
  1. I know that a banana is called *banana*
  2. If the speaker meant to refer to the banana, she would have asked me to show her the banana
  3. But she didn’t, she used a strange word, *fendle*
  4. So she must intend to refer to something other than the banana
  5. A plausible candidate is the whisk.
  6. *Fendle* must refer to the whisk.
• ME can also be explained in terms of Clark’s principle of contrast—a pragmatic principle stating that every difference in form corresponds to some difference in meaning. Avoid synonyms.
• On a more general pragmatic account, these cases are explained by the listener assuming that the speaker names things that are new to the discourse context

Golinkoff et al. (2000)
• Lexical principles: strategies that emerge in the course of cognitive and lexical development, and constrain the problem space for word learning.
• Principles start out in an immature form.
• There is differential sensitivity to different cues at different stages of development:
  - Perceptual saliency cues and associative learning
  - Social cues and referential intent+linguistic cues for mature principles

Ongoing debate
• a) constraints or principles are not specific to words (e.g. can also apply to facts as referential acts, e.g. Markson and Bloom, 1997) and
• b) can be explained by application of more general pragmatic inferences on the speaker’s referential or communicative intent (e.g. Clark, 1987; Akhtar et al. 1996; Tomasello et al., 1996; Bloom, 2000; Diesendruck and Markson, 2001).

Shape bias and substance bias (Soja, Carey and Spelke, 1991)
• When objects are solid (e.g., a hammer, a pencil, etc.), children tend to categorize them by their shape.
• When objects are not solid (e.g., hair gel, liquids), children tend to categorize them by their substance.
Note: in some languages, these distinctions are expressed in the language (mass nouns vs. count nouns in English), but in other languages, these distinctions are not expressed (e.g., classifier languages, such as Japanese). However, children of both languages respond similarly, which shows that this is not a language effect, but a cognitive effect.

Constraints theory
• Good points:
  – explains speed at which new words learnt (8-10 words per day during 1st year, Carey, 1978)
  – Upheld in Markman’s lab (Markman, 1989, 1992, 1993) in studies on 3-5 year olds
• Bad points:
  – how do children decide whether to apply object or action constraint?
  – Why are so many first words parts of objects (leg, eye, head)
  – how do children learn one object may have more than one name (e.g. dog, Rover, animal)

Syntactic bootstrapping
• Children build on syntactic cues in constraining potential meanings
• At least regarding the syntactic category of the word: e.g. if something has a verbal inflection, it is a verb
• Experiments: children at 17 months can already distinguish between count and mass nouns based on syntax (this is a sib versus this is sib.)
• Verb meanings are also supported by syntax: transitive and intransitive meanings are
distinguished on the basis of structure (pl. Fisher et al. 1994)
• Language context undoubtedly helps, but this maybe specific to function words, not to abstract structure especially at 2 years of age.
• 3-4 year olds already know enough syntax for abstract structure to help them in guessing word meaning
• E.g., video with a man poring some substance from a jar into a bowl (Brown 1957)
  Point at: ‘daxing’ (poring) ‘some dax’ (substance) ‘the dax’ (jar or bowl).

Social pragmatic cues
• Children are effective word learners because they have the ability to infer the others communicative intent that is relevant from the point of view of the given situation
• The ability to follow the speaker’s direction of gaze is a precursor of both language and theory of mind or intention attribution

Relying on speaker’s direction of gaze in referential understanding
(Baron-Cohen, Baldwin, Crawson, 1997)

• Baldwin 1991, 1993: the key strategy normal children use in solving the problem of referential indeterminacy:
  • Speaker’s direction of gaze (SDG) strategy: assuming that a novel word uttered by the speaker refers to the object the speaker is looking at
  • Alternative:
  • Listener’s direction of gaze (LDG) strategy: assuming that a novel word uttered by the speaker refers to the object the listener is looking at
  • Test situations: follow-in labeling and discrepant labeling

Tomasello és Barton, 1994
‘Let’s find the Toma’
Social-pragmatic cues
• Tomasello et al: not only direction of gaze, posture and other built-in cues; flexible adaptation to the situation in understanding social-pragmatic cues
• Not a simple association
• More domain-general mechanism
• Language acquisition is a form of cultural learning

5. LANGUAGE ACQUISITION II

-ACQUISITION OF GRAMMAR
   Holophrastic (one part of speech)
   Pivot-Open (two parts of speech)
   Telegraphic (four parts of speech)
   Adult (eight parts of speech)
   Linguist (each part of speech has many sub-categories)

-ACQUISITION OF MORPHOLOGY
  • Once acquired, morphemes can be used in a very productive way (Berko, 1958).
  • more salient morphemes are learned first. E.g., progressief -ing is learned before third person sg. -s. The former is phon. more salient, and contributes more to the semantics of the whole-word form.
  • U-shaped learning of morphology:
    From 1. Holophrastic: men, went, broke, brought
    To 2. overgeneralization (e.g., bringed, goed, etc.)
    To 3. Knowledge of both Rules and Exceptions to the Rules: men, went, broke, brought
  • overgeneralization applies to derivational rules as well:
    • ‘you have to scale (weigh) it’ ‘he’s keying (opening) the door’
    • Morphological awareness is great source for language creativity
    • papegaai ‘parrot’ => papa gaai en äiti gaai; loopvliegen ‘walkflying’
  • sometimes, the multiple possibilities to express a certain meaning with a morphological affix might generate some problems!

5.1. Grammar: two word stage

The two-word stage is also called the Pivot-Open stage because one of the words is usually a Lexical Word (an open set that refers to something), and the other word is a Functional Word (a closed set with grammatical rather than reference meaning).

Allgone sock. Pretty boat.
Byebye boat. Pretty fan.
More wet. More taxi.
Katherine Sock. More melon.
Hi Mommy. Push it.
Allgone sticky. Move it.
It ball. Mommy sleep.
Dirty sock. Bye-bye melon.
See boy Bye-bye hot. (Adam, Eve, and Sarah)
See soci. (Fromkin Rodman Hyams [2011] 369-370)
Telegraphic Speech
Functional categories like Determiners, Auxiliaries, Prepositions, Conjunctions and Expletives are missing.
And the Lexical categories like Nouns, Verbs, Adjectives, and Adverbs (usually without any suffixes) are present.
Cat stand up table.
What that?
He play little tune.
Andrew want that.
Cathy build house.
No sit there.
(Fromkin Rodman Hyams [2011] 347)

Acquisition of grammatical morphemes (Brown, 1973).
• 14 grammatical morphemes
• spontaneous speech of three children
• Acquisition criterion of 90% correct use in obligatory contexts (interval of 16 months from 0 till 100%).
  – 1. Order of acquisition
  – 2. Grammatical complexity
  – 3. Semantic complexity
  – 4. Frequency in the input
• Order of acquisition correlates with semantic complexity, but not with frequency in the input.

Acquisition of grammatical rules
• Do children learn a language just by imitating, or are they making hypotheses about underlying rules?
• Word combination shows systematicity already
• Overgeneralization errors
• 1. Experimental tests of productivity (The WUG test)
• 2. Irregular past tense is learned in three stages (U-shaped curve)

Word Inflections
• Function word sequences:

<table>
<thead>
<tr>
<th>Function</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plurals:</td>
<td></td>
</tr>
<tr>
<td>1. All singular</td>
<td>1. -ing</td>
</tr>
<tr>
<td>2. Some irregulars</td>
<td>2. Plural –s</td>
</tr>
<tr>
<td>3. Regular ‘s’ overgeneralized</td>
<td>3. Possessive –’s</td>
</tr>
<tr>
<td>4. [-ez] for all</td>
<td>4. 3rd person singular –s</td>
</tr>
<tr>
<td>5. Only irregulars remain problematic</td>
<td>5. Past marker –ed</td>
</tr>
<tr>
<td>6. Irregulars memorized</td>
<td>6. Future marker ‘will’</td>
</tr>
<tr>
<td>7. Verb ‘to be’ (is, are)</td>
<td></td>
</tr>
<tr>
<td>Early combinations:</td>
<td></td>
</tr>
<tr>
<td>agent + action</td>
<td>‘Daddy sit’</td>
</tr>
<tr>
<td>action + object</td>
<td>‘drive car’</td>
</tr>
<tr>
<td>agent + object</td>
<td>‘Mommy sock’</td>
</tr>
<tr>
<td>action + location</td>
<td>‘sit chair’</td>
</tr>
</tbody>
</table>

Acquisition of grammatical morphemes (Brown, 1973).
• 14 grammatical morphemes
• spontaneous speech of three children
• Acquisition criterion of 90% correct use in obligatory contexts (interval of 16 months from 0 till 100%).
  – 1. Order of acquisition
  – 2. Grammatical complexity
  – 3. Semantic complexity
  – 4. Frequency in the input
• Order of acquisition correlates with semantic complexity, but not with frequency in the input.
• entity + location  ‘toy floor’
• possessor + owned  ‘my teddy’
• entity + attribute  ‘crayon big’
• Demonstr. + entity  ‘this phone’

5.2. Acquisition of Morphology

AGE 2:
Progressive –ing: I am singing.
Plural –s: blue shoes.
Copula am, is, are: He is asleep.
Articles a, the: He is a doctor.
(Aitchison 574)

AGE 3:
Third Person Singular –s: He wants an apple
Past tense –d: I helped Mummy
Full Progressive be + -ing: I am singing
Shortened Copula: He’s a doctor
Shortened Progressive: I’m singing

Lack of negative evidence?
Brown and Hanlon,
CHILD: Nobody don’t like me.
MOTHER: No, say “Nobody likes me.”
CHILD: Nobody don’t like me.

(dialogue repeated eight times)

MOTHER: Now, listen carefully, say “Nobody likes me.”
CHILD: Oh, nobody don’t likes me.
(Fromkin Rodman Hyams [2011] 326)

Feedback on pronunciation and truth values, not on syntactic well formedness. Feedback on form is ineffective (in the above example, even after 8x). A problem for acquisition theories building on reinforcement and imitation.

Imitation?

CHILD: My teacher holded the baby rabbits and we patted them.
ADULT: Did you say your teacher held the baby rabbits?
CHILD: Yes
ADULT: What did you say she did?
CHILD: she holded the baby rabbits and we patted them.
ADULT: Did you say she held them tightly?
CHILD: o, she held them loosely
(Cazden 1972)

CHILD: Want other one spoon, Daddy.
FATHER: You mean, you want the other spoon.
CHILD: Yes, I want the other one spoon, please Daddy.
FATHER: Can you say, “the other spoon”?
CHILD: Other ... one ... spoon.
FATHER: Say “other.”
CHILD: Other.
FATHER: Spoon
CHILD: Spoon
FATHER: Other spoon.
CHILD: Other ... spoon. Now give me other one spoon?

Overgeneralizations: argument structure
Don’t giggle me.
I danced the clown.
Yawny Baby—you can push her mouth open to drink her.
Who deaded my kitty cat?
Are you gonna nice yourself?

CF: Colorless green ideas sleep furiously.
(Fromkin Rodman Hyams [2007] 361)

Acquisition of Negatives
Stage One: “No you catch me.”
Stage Two: “You didn’t caught me.”
Stage Three: “You didn’t catch me.”

• 1st stage - attach no/not to beginning of sentence (sometimes at end)
• 2nd stage – negatives appear between subject and verb (don’t stayed at beginning in imperatives, but not can’t)
• 3rd stage – appearance of nobody/nothing & anybody/anything & inconsistent use of “to be” verb is and auxiliary “dummy” do verb.

Question Formations
• 1st stage – wh- word placed in front of rest of sentence: Where daddy go? What Mummy doing? Why you singing?
• 2nd stage – addition of an auxiliary verb: Where you will go?
  Why kitty can’t see? Why you don’t know?
• 3rd stage – subject noun changes places with the auxiliary: Where will you go? Where will you go? Why can’t kitty see? Why don’t you know?

Comprehension of syntax
• Command of certain properties of adult language even in initial stage of grammatical competence, e.g., word order. Videos: Big Bird is tickling Cookie Monster & Cookie Monster is tickling Big Bird. Upon sentence presentation, 17 months old children tend to watch the right video.
• Transitive/intransitive distinction. Again two video’s and two sentences were presented to 27 months old children. 1. Big Bird is gorping with Cookie Monster and 2. Big Bird is gorping Cookie Monster. In 1st case, children tend to watch to video where both animals perform action, in 2nd case to video where Big Bird is doing something to Cookie Monster.
5.3. How do we acquire language?

Task: combining and modifying words **correctly** and **productively**

- Two theoretical approaches
  - Nativist: children are born with INNATE KNOWLEDGE which helps them in making sense of the world around them
  - Constructivist/emergentist/empiricist: children LEARN language, by constructing linguistic knowledge based on input from the environment

**Argument for innateness: Language Universals**

What evidence is there for innate knowledge of certain basic language features present in all human languages?

LINGUISTIC UNIVERSALS > UNIVERSAL GRAMMAR

All languages have:
- A grammar
- Basic word order (in terms of SOV, etc.)
- Nouns and verbs
- Subjects and objects
- Consonants and vowels
- Absolute and implicational tendencies
  E.g., if a language has VO order, then modifiers tend to follow the head

**Language universals**

- Chomsky: there are certain universal constraints on rules and categories
- These constraints are biological and so innate
- Language acquisition device, Universal Grammar, that accounts for all possible human languages
- Principles and parameters

**Language universals**

- Substantive universals
  - Categories necessary for analysing or constructing languages: noun, vowel, subject, etc.
- Formal universals
  - Constraints on the types of linguistic rules
- Implicational universals: Of the form if x, then y
  - If a language marks gender on nouns, it will also mark it on pronouns.
  - If a language is predominantly VSO in its word order, then the adjective will most often follow the noun.

**“Universal Grammar”**

Humans then learn to specialize this “universal grammar” (UG) for the particulars of their language.
- Word order, syntactic rule preferences
- Phonetic and phonological constraints
- Lexicon
- Semantic interpretations
- Pragmatic ways to converse
Innate behaviors: Lenneberg’s criteria
1. Maturationally controlled, emerging before they are critically needed
2. Do not appear as the result of a conscious decision.
3. Do not appear due to a trigger from external events.
4. Are relatively unaffected by direct teaching and intensive practice.
5. Follow a regular sequence of “milestones” in their development.
6. Generally observe a critical period for their acquisition

Argument for innateness:
Universal developmental pathway
• In spite of different backgrounds, different locations, and different upbringings, most children follow the very same milestones in acquiring language.
• Is this criterion met?

Argument for innateness: critical period
What is a critical period?
• For first language acquisition, there seems to be a critical period of the first five years, during which children must be exposed to rich input. There is also a period, from about 10-16 years, when acquisition is possible, but not native-like.
• For SLA, the issue is more complicated... More on that later.
• Is this criterion met?

The Critical Period Hypothesis
• CPH: Proposed by Lenneberg
  – There is only a small window of time for a first language to be natively acquired.
  – If a child is denied language input, she will not acquire language
• Genie: a girl discovered at age 13 who had not acquired her L1 (-- Isabelle and Victor)
• Normal hearing child born to deaf parents, heard language only on TV, did not acquire English L1
  – Second Language Acquisition:
    • Younger learners native fluency.
    • Older learners (>17) never quite make it.
  – Aphasia:
    • Less chance of recovery of linguistic function after age 5.

Nativism
• We need to postulate the existence of an innate grammar since
  – Linguistic input is not sufficient and not appropriate for the acquisition of grammar
    (poverty of the stimulus argument)
• The speech children hear is full of errors
• Grammar is too complex, it is impossible to get to it based on heard examples only
• Children say and understand words that they never heard or are ungrammatical in an adult grammar. Errors are systematic.
  – Discovering the grammar would require corrections that adults do not give and children are not able to exploit. (lack of negative evidence argument)
• Main proponents, rationalists/nativists, Chomsky and Pinker.
- Language not learned, but grows, assisted by LAD
  - \( \Rightarrow \) Universal Grammar, parameter setting

- Nature: specific idea: parameter setting
- Restricted variability across languages
- Role of language experience is to set parameters for language being acquired
- Example: parameter for obligatory pronoun use:

  - Children construct rules which they then use creatively
  - Nobody teaches them these rules explicitly; adults do not have an explicit knowledge of phonological, morphological etc. rules. Language acquisition proceeds very quickly in the absence of formal education.
  - Children are not able to repeat sentences that their grammar is not yet able to generate
    
    **Mom:** He is going out. **Child:** He go out.
    **Mom:** That’s an old-time train. **Child:** Old time train

**Continuity assumption**
- Babies have the same abstract grammatical knowledge as adults
- How do they map it onto specific linguistic elements?
- Semantic bootstrapping (Pinker, 1984):
  - Acquisition of simple meanings first
    - Innate canonical linking rules: mappings that are only partially valid, but give fairly good mappings between formal (subject, noun, etc) and conceptual (agent, inanimate object) categories, which are then revised by learning.
    - Syntactic function of words is inferred based on their meaning
    - They are not universal and not bidirectional, but agents of an event tend to be subjects of sentences, verbs generally denote actions. Not all nouns are object names, but object names are nouns

**Problems with the semantic bootstrapping theory**
- Not supported by data
- Slobin (1985) crosslinguistic and diachronic analyses: there are no linking rules that would work well for all languages
- Many contradictory utterances in early child language:
  (e.g. subject is not an agent) I like .., I see .., my tummy hurts, It has buttons).
- Builds on conceptual bases in the beginning, but linking rules are not determined by development, but innate constraints of universal grammar.
Slobin’s operating principles and universals (1973)
Whether parts of language acquisition are innate or not, developing kids seem to follow specific strategies and their acquisition processes reveal universals

**Operating Principles:**
- A. Identify word units.
- B. Forms of words may be systematically modified.
- C. Pay attention to the ends of words.
- D. There are elements which encode relations between words.
  - postposed forms learned before preposed forms
  - articles before nouns less salient than noun suffixes
- E. Avoid exceptions

**Universal 1:**
- one-to-one marking is acquired earlier than
- compound markings.
- unchanging singular articles like French le are
- acquired faster than der/den/dem in German

**Universal 2:** Stages of linguistic marking
1. no marking: bird, birdy - singular & plural
2. appropriate marking in limited cases: bird, birdies – plural
3. overgeneralization of marking: mens, sheeps
4. appropriate marking everywhere

**Universal 3:**
- if a group of inflections all mark the same relation, the child will tend to use the single commonest form for all cases
- irregular past tenses reduce to dental suffix –ed

**Universal 5:**
- semantically consistent grammatical rules are acquired early and without significant error
- kids don’t overextend the progressive to stative verbs, as in *She’s knowing the answer* or *I'm liking cookies*

**Constructivist approach (Tomasello 2000)**
- Babies don’t have abstract knowledge of grammar
- Simpler and more general learning mechanisms
- Neither innate representations nor special acquisition devices or processes are necessary
- Poverty of stimulus is not valid
- Children’s early utterances follow adult utterances heard by them quite closely

**Simple learning mechanisms (Tomasello, 2000)**
- Children start with semantic categories based on general cognition (agent, action).
  These are then gradually extended when they meet new words in the „agent“ slot, until the category develops into the category NOUN (Braine 1992)
- Distributional analysis (e.g. words that can have –s as an ending can also have –ed and –ing, these have to be verbs, Maratsos & Chalkley (1980))
- categorization
- Intention reading
• Analogy, structure mapping, generalization
• Errors on this approach are errors of analogical extension

**Constructivism**
• Grammatical rules are extracted from the input using general learning mechanisms.
• Input constrains different potential sources of information and makes it possible for the child to discover grammatical categories that way
• Phonological bootstrapping
  – *rEcord* noun *recOrd* verb
• Prosodic bootstrapping
  – Motherese exaggerates prosody
• Children do get indirect negative feedback. Repetitions and extensions are related to grammatical errors
  – Adults are more likely to repeat well-formed sentences
  – They are more likely to ask for clarification after incorrect sentences
• Development in other areas predicts grammatical development
  – Vocabulary at two predicts grammar at two and a half years (Bates et al. 1988)
  – The ability to combine actions during play correlates with the ability to combine words (Shore et al., 1984)
6. VISUAL WORD RECOGNITION

6.1. Word recognition

Skilled readers ca. 300 words per minute \( \rightarrow \) ca. 200 ms recognition of one word

**Word recognition** =
- we identify the word as familiar
- We access information about the word: meaning, lexical category, pronunciation, etc.
- Context and the aim of reading has an effect on how much information becomes available (we can recognize it as a word, but do not really get the meaning)

**Lexical processing**
- *the means by which single words are recognized*
- People do it very fast
- Average person has around 50,000 words in memory
- Takes only 250 msec to find a word from among 50,000
- It is automatic and robust
- Word recognition is a retrieval task, not compositional like sentence processing
- The auditory or visual signal must be mapped onto representations of known words in the listener’s mental lexicon
- Sounds relatively simple, but words are not highly distinctive (tens of thousands of words are constructed from 30 to 40 phonemes)

**A simple model**
1. Processing visual features
2. Combining features into letters
3. Combining letters into words

**Features and context**

<table>
<thead>
<tr>
<th>Why visual features are important</th>
<th>Why context is important</th>
</tr>
</thead>
<tbody>
<tr>
<td>EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE</td>
<td>Word superiority effect Reichle 1969</td>
</tr>
<tr>
<td>EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE</td>
<td></td>
</tr>
<tr>
<td>EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE</td>
<td></td>
</tr>
<tr>
<td>EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE</td>
<td></td>
</tr>
</tbody>
</table>

---

Psycholinguistik – Folien, WS11
**Differences between written and spoken words**

- Spoken word recognition takes place in time - words are not heard all at once but from beginning to end.
  - Written words are available to readers as a whole (depending on length).
- Typically there is no chance to reconsider the spoken input.
  - In printed text we typically can re-read words or passages.
- Spoken words are rarely heard in isolation but rather within longer utterances, but there is no reliable cue in speech to mark word boundaries.
  - In printed text white space unambiguously mark word beginnings.
- In spoken words phonemes are realized differently in different contexts (coarticulation). “Sweet girl” is often pronounced as “sweek girl” (but also within syllables → compare tongue position with /ki/ versus /ku/).
  - Usually no such variability is found in printed letters.

**Word recognition**

- What makes it difficult?
- The time required to recognize a word can reflect the level of difficulty
- Measures of recognition are usually indirect

**Indices of recognition**

- Fixation times during reading (eye-tracking)
- Naming latency
- Lexical decision times
- Categorization times
- Gating
- Tachistoscopic recognition

**Factors affecting word recognition**

- Perceptual unambiguity
- Length
- Frequency
  - To be controlled for
  - Interaction with length
  - Methodological problems (written-spoken, individual variability)
  - Frequent words are accessed faster and more easily than less frequent ones.
    Effect is task-dependent. Balota & Chombley (1984): strong frequency effect in lexical decision, moderate in naming and only minor effect in category decisions. The authors claim that post-access processes are responsible for frequency effects.
- Age of acquisition
  - Frequency effects are in fact ~ effects?
- Uniqueness point
  - Early uniqueness point = *strawberry* (there are no other English words beginning with /str/)
  - Late uniqueness point = *blackberry* (not unique at /b/ of *berry*: *blackbird, blackbeetle, ...*)
- What affects lexical access time?
  - Faster responses to words with earlier uniqueness points
  - E.g. *strawberry* vs. *blackberry*
  - Again, this effect does not necessarily tell us anything about the organization of the mental lexicon.
• Many of these factors target the ease of mapping the input onto a unique lexical representation
  – Uniqueness point
  – Word length
  – Neighborhood
• Others reflect the ease of accessing the lexical representation stored in memory
  – Frequency
  – Phonological and semantic priming
  – Concreteness
• neighborhood effects
  – Neighborhood size refers to number of words you can create by changing one letter (rind, find, kind, mind, bind)
  – Interaction with frequency: Low F words with large N are recognized faster than Low F words from small Ns (Grainger, 1990; rind vs. ruin)
• regularity
• Imageability/concreteness: to what extent we are able to form a mental picture of the word (apple is easier to access than freedom)
• syntactic category (noun vs. verb)
• morphological complexity

**Primting**
• Repetition priming:
  – Effect can last for hours even
  – Bigger effect on rare words
  – Debate: ~ due to activation of stored representation of the item or the storage of the processing event in episodic memory
  – Support for the latter: repetition priming usually observed within modality, but semantic priming works cross-modally as well
• Phonological
  – Hard to demonstrate
  – Only works with short SOA
  – The specific shape of items is important, it can be inhibitory as well
  – Alphabetical ordering or rhymes?
• Associative
• Semantic
• Morphological
• Attentional

6.2. **Reading studies**

**Measures:**
– Position of fixations
– Fixation time
– Sequence of fixations

**Processing:**
– Foveal: 2 degrees around fixation point (1 degree cca. 4 characters)
– Parafoveal: 5-5 degrees on either side of fixation point
– Periphery: everything outside the above
**Eye movements**

- Types
  - Fixations
  - Saccades
- What is their respective purpose?
  - Fixation: acquiring new information
  - Saccade: movement of the eye towards a new point of attention
- Differences in eye-movements between reading and listening comprehension?
  - Reading: Eye-movements on words
- Types of measurements in reading?
  - First pass: Duration of all fixations in a region before leaving it to the right or left.
  - Second pass: Duration of the fixations in a region after having left it either to the left or right
  - Regression Path Duration: The sum of the durations of fixations in a region, plus the fixations from that region to earlier parts of the text until the region is left in a forward direction
- # Total time: The sum of all fixations in a region

**Questions in eye movement studies**

- What kind of information is extracted (feature, orthographical, phonological, semantic)
- When and how this information is accessed (parafoveally, foveally, before/after launching the saccade program)
- How we make decisions about when and where eye movements should be targeted

**Advantages of eye movement studies**

1. Eye movements are a part of normal reading, participants don’t have to do anything that would disturb the reading process
2. The processing of a word in context is reflected in fixation time.
3. As fixations are relatively short (200-250 msec), it sets a time limit for processing.
   - This is the time available for lexical processing

**Eye movements in reading**

- These eye movements are not smooth; short, cca. 200 msec fixations, which are separated by saccades and regressive fixations
- During every fixation, we extract information cca. 4 characters to the left and cca. 14 characters to the right
  - *The quick brown fox jumped over the lazy dog.*
- Then saccade cca. 8 characters to the right; fixations overlap
- 10-15% of fixations are regressive
- We do not read every word
  - Fixations on cca. 80% of content words
  - Much less on function words, we skip short words easily, we hardly never skip words of 8 or more characters
- Carpenter and Just (1981) Fixation length bigger for longer and smaller for more frequent words
Lexical processing
- Verbs longer than nouns
- Ambiguous words longer
- Context has an effect (priming)

Naming
- Participants have to read lists of words and nonwords.
- The pronounce words faster than nonwords, because processing in this case is speeded up by lexical access, and frequent words are read out faster than rare words

6.3. Models of visual word recognition
- Serial search models
  - Forster’s autonomous search model
- Parallel detection models
  - Morton’s logogen model — different variations
  - Connectionist models
- Interactive activation model
- Mixed models
  - Becker’s verificational model
Serial versus parallel models
• serial search model – analogy of search in a paper lexicon
  – The lexicon is autonomous, i.e. independent of other systems involved in language processing.
  – activation of words is not directly influenced by syntactic or semantic factors. Information from the lexicon is fed into this more general system, and in this way syntactic/semantic information may influence word activation.
  – the word recognition system has two components: the orthographic properties of the word; the phonetic properties of the word.
  – The lexicon is organized in descending order of frequency, i.e. more frequent words are searched before lower frequent words.
• parallel/direct access models – perceptual properties of the word directly access the word’s lexical entry (through activating its detector), more than one lexical entry is activated at a time, in parallel.

Search in a paper lexicon
speculum (‘spTkjTtTm) n., pl. -la (-lT) or -lums. 1. a mirror, esp. one made of polished metal......
sped (spTd) vb. a past tense and past participle of speed.
speech (spi:t*) n. 1. the act or faculty of speaking.....
speechless (’spi:t*ls) adj. 1. not able to speak. 2. temporarily deprived of speech. 3. not expressed.......
speed (spi:d) n. 1. The act or quality of acting or moving fast; rapidity. 2. the rate......

Forster (1976) autonomous search model
• Serial search
• Library metaphor: a word (book) can be found only at one place in the mental lexicon (library), but it can be found through different sources:
  – Words are searched in one of the 3 big access files: orthographic – based on visual features (for visual word recognition)
  – phonological – sound (auditory word forms)
  – syntactic/semantic – meaning and grammatical category (for production)

**Changes in serial models**
• Becker (1979): search can be based on semantic features, helped by connections among words that are semantically related, these are activated, accounts for priming effect.
• Glanzer and Ehrenreich (1979) two lexicons: a big, full lexicon, containing every word, and a shorter, containing only very frequent words.
Marslen-Wilson’ Cohort model
- was designed specifically to account for auditory word recognition
- start of words activates all words that begin with that sound. This set of words is called a cohort. (bottom-up)
- Once initial cohort has been activated, further information (both bottom-up and top-down) serves to narrow the choices
- You recognize a word when you are left with a single choice. This is the point at which it diverges from other words – uniqueness point

frequency effects greater experience results in higher baseline activation levels for frequent words

**semantic priming**: Two-way information flow between cognitive system and logogen system, activation in one logogen (indirectly) spreads to all related words, these are primed.

The new logogen model

- Morton, 1979, did not find any crossmodal priming effects → changed the model, information enters separate logogen systems.

love

individual thresholds

loathe

[Diagram of the new logogen model]

- Visual logogen
- Auditory logogen
- Target logogen
- Visual analysis
- Auditory analysis
- Response buffer
1. Access stage (prelex)
2. Selection stage (prelex)
3. Integration stage (postlex; it is only here that activation for alternative candidates fades)
Parallel, interactive, direct access

Marslen-Wilson’s Cohort model, 1989
- In the first version of the model, context had a very strong influence
- The first version could not explain how we can identify a word with a wrong beginning (sigaretta), or hearing only its second half
- Marslen-Wilson 1989:
  - bottom-up priority, context cannot constrain initial cohort (like before), and can only have an effect beyond the uniqueness point, in the integration stage
  - Exclusion of a candidates is not all or nothing: competition raises one candidate above the others, but other candidates are not deleted, only partially deactivated (makes the model very similar to the TRACE model)
- Items with the highest rank are accessed parallel for comprehension — the item with the best fit is integrated and becomes recognized

**Frequency effects in word recognition**
- Auditory lexical decision task, with words of the same length and same uniqueness point, different frequencies
e.g.
  DIFFIC | ULT frequent (250ms)
  DIFFID | ENT infrequent (379ms)
- The original cohort model could not account for this difference

**Zwitserlood’s experiment**
- Crossmodal priming

Acoustic prime: Captain

Visual test word: slave ship shop

<table>
<thead>
<tr>
<th>Input</th>
<th>Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>All s- words: sea, salt, select, saw, spinach, etc.</td>
</tr>
<tr>
<td>sp</td>
<td>All sp- words: spirit, spinach, spoon, speculate, etc.</td>
</tr>
<tr>
<td>spe</td>
<td>Spectrum, speculate, spell, special</td>
</tr>
<tr>
<td>spec</td>
<td>Spectrum, special, speculate</td>
</tr>
<tr>
<td>spect</td>
<td>Spectrum</td>
</tr>
</tbody>
</table>
• Priming to both alternative only in the „early” condition: at this point, both words are listed in the cohort
• Bigger priming to “ship” — frequency effect
• Word frequency influences activation level of candidates in early stages of lexical access

Connectionist models
McClelland and Rumelhart, 1981, 1982, interactive activation model

Interactive Activation model
• To explain Reicher-Wheeler’s word superiority effect.
• Top-down (word, letter) excitatory connections help letter identification in words, but not in nonwords
• Localist – features, letters and words are represented by single nodes—their representation is not distributed in the network

TRACE model
• Like the interactive-activation model of printed word recognition, TRACE has three sets of interconnected detectors
  – Feature detectors
  – Phoneme detectors
  – Word detectors
• These detectors span different stretches of the input (feature detectors span small parts, word detectors span larger parts)
• The input is divided into “time slices” which are processed sequentially.
• Within a set (or level) connections are inhibitory
  – E.g. evidence that a certain stretch of the input is the word “tip” is evidence that it is NOT any other word
• Between sets (or levels) connections are excitatory
  – E.g. evidence that a certain stretch of the input is the sound /t/ is evidence that it might be the beginning of the word “tip”
  – Also, evidence that the word is “tip” is evidence that its parts are /t/ /i/ /p/, so there are “top-down” (feedback) effects in TRACE as in the interactive activation model
• Or inhibitory
  – If it’s a /t/ it isn’t the beginning of “cat”

• Accounts for context effects
• Can handle (some) acoustic variability (and noise)
• Can account for phoneme restoration (Warren - “Open the oor” heard as “Open the door”)
• Can account for co-articulation effects
• Can find word boundaries (using the possible word constraint)

### Evaluation of models
• All of the models provide and explanation for the word frequency effect
• Logogen: each time a word is encountered, the threshold for that logogen is temporarily lowered. With high-frequency words, the recovery from the lowering of the threshold is less complete than with low-frequency words, so less sensory information is needed for recognition.
• Search: Frequency effects are explained in terms of how words are stored in the various files.
• Cohort: Many word candidates are activated in the initial access phase, but more frequent words would be chosen in the subsequent stage.
• All of the models can account for semantic priming
• Logogen: there is a rapid and temporary lowering of the threshold of those logogens that are related to a prime.
• Cohort: the prime narrows the set of candidates in the initial cohort list; a shorter initial cohort leads to faster recognition of a target word.
• Search: with each word, we generate a list of words that might come next. Priming is conceived of more as a controlled than an automatic process.
• Cohort model is more explicit about the time course of spoken word recognition and thus is better able to explain how sounds in different positions within the word may affect recognition.

6.4. Lexical ambiguity

• Do ambiguous words (bank) have more than one representation in the lexicon?
• As readers or listeners we don’t usually notice the ambiguity, but what effect does it have?
  Do we consider multiple meanings of ambiguous words when we hear or see one?
• How might the sentence context influence how lexically ambiguous words are processed?
  The appropriate meaning is usually determined by context

• My Blackberry Is Not Working! - The One Ronnie, Preview - BBC One
• http://www.youtube.com/watch?v=kAG39jKi0lI&feature=player_embedded

Lexical Ambiguity
• Hogaboam and Perfetti (1975)
  – 1. The jealous husband read the letter
  – 2. The antique typewriter was missing a letter
• (1) requires the activation of the dominant postal meaning;
• (2) requires activation of the secondary sense.
• Participants were given a series of sentences such as these and were asked to decide whether the final word in the sentence was ambiguous
• Decision times were faster when the sentence required the secondary sense than when it required the primary meaning.
• Explanation: The common meaning is easily activated, so the time taken to find the other meaning is more directly related to response times in this task.

Lexical Ambiguity - MacKay 1966
• Sentence completion task
  – After taking the right/left turn at the intersection, I....
• Harder to complete after “right” (ambiguous) than “left” (unambiguous)

Lexical Ambiguity - Models
• Context-guided (selective) access (Schvaneveldt)
  – Appropriate meaning is chosen by context, others are not considered
  – But how could it work?
• Ordered Access (Hogaboam & Perfetti, 1975)
  – Most common meaning checked first
  – Accepted if it fits context
  – Otherwise other meanings are checked
• Multiple Access (Swinney, 1979)
  – All meanings accessed, context selects among them
• Reordered Access Model (Duffy, Morris, & Rayner, 1988)
Contradictory evidence from the 1970s

- For selective access: e.g. Hogaboam & Perfetti, 1975,
  ambiguity detection task: Found longer RT when word used with its more common meaning (e.g. ink pen, rather than sheep pen)
  - The accountant filled his pen with ink.
  - The farmer put the sheep in the pen.

- For multiple access: e.g. Foss, 1970,
  phoneme monitoring: People are slower to detect /b/ in A) than in B) because “straw” is ambiguous, even though the context (“farmer”) strongly favours one meaning. Suggests both meanings accessed. (compare the old MacKay finding)
  - A) The farmer put his straw beside the machine.
  - B) The farmer put his hay beside the machine.

  However, this task is sensitive to the length of preceding words. A short word may not be fully processed before the next word begins. A longer word can be identified before their end. Problem: majority of polysemous English words are short. If this is controlled for, the effect disappears (Mehler et al., 1978).

- So which view is correct?

Swinney, 1979
Context: none (bugs or insects) or biasing (spiders, roaches, and other bugs/insects)
Words: Ambiguous (bug) or unambiguous (insect)

Rumour had it that for many years, the government building had been plagued with problems. The man was not surprised when he found several (spiders, roaches, and other) bugs (insects) in the corner of his room.

```
ant  sew  spy
```

In both contexts + ambiguous

1. Ant = spy, > sew
2. Ant > spy and sew

Replicated by........
- Onifer & Swinney, 1981 with biased ambiguous words;
- Tanenhaus, et al., 1979, naming task - context-independent meaning fades after _ 200 msec;
- Seidenberg, et al., 1982, for a few 100 msec all meanings of an ambiguous word are activated regardless of semantic and syntactic constraints.
- Results support a modular view of sentence processing. Challenges by some, but generally accepted view that lexical ambiguity is resolved by an interaction b/w frequency and context
Lexical Ambiguity - Balanced and Biased Ambiguities
• The meanings of some ambiguous words (balanced) are roughly equally common
  – For others (biased) one meaning is much more common than the other(s)
• Onifer & Swinney (1981) replicated Swinney’s (1979) results for biased ambiguities
• However, others have claimed that only balanced ambiguities show multiple access

Lexical Ambiguity - Types of Context
• Contexts may be more or less strongly biased towards one or other meaning of an ambiguous word.
  – Maybe selective access occurs only with strongly biasing contexts
• Contexts may be consistent with specific properties of one or other meaning of an ambiguous word (Tabossi)

Lexical Ambiguity: Conclusions
• The Swinney 1979 study provided striking evidence for multiple access
• Later studies have found that evidence for multiple access is clearest with balanced ambiguities and contexts that are not too constraining
7. MORPHOLOGICAL PROCESSING

- **Morpheme**: smallest unit with meaning
- Its actual realization is a morph; if it has several phonological realizations, these are called allomorphs of the same morpheme:
  - /t/ and –/d/ and /id/ are three allomorphs of the same –ed morpheme
- **Free and bound morphemes**
  - Bound morphemes: e.g. verbal prefixes (as pre- in preheat) and suffixes
- **Derivational morphemes** (-ment in development) and **inflectional morphemes** (-s in develops).

Morphology is important in language processing

Intuition: following morphological principles we make new words every day and comprehend new complex words

**Observations:**
- Speech errors are morphologically structured: it’s not only we who have screw loose (for “screws loose”); easy enough (for “easily enough”).
- Affix perseverations: ministers in the churches
- Children’s overgeneralizations: *goed* (for went)
- Patients with aphasia often produce neologisms that have morphological structure: he *jumberfoked* and off he went

- **Holistic storage**: every word has its own access representation: //
  - /walk/ /walked/ /walking/...
  - We access morphologically complex words holistically
- **decomposition**:
  - only morphemes (stems and affixes) have their own access representations: /pre/ /develop/ /ment/ /al/....
  - Morphologically complex words are analyzed into their building blocks (prelexical analysis)
- The entire word form and its morphemes all have their access representation: /predevelopmental/ /pre/ /develop/ /ment/ /al//developmental/...
  - Morphologically complex words can be analysed or accessed holistically
  - When do you need decomposition?
- cognitive economy?
- What are the influencing factors?

7.1. Decompositional or holistic (ganzheitlich)?

- May depend on experimental design. Rubin, Becker and Freeman (1979) found a processing difference between suffixed and pseudosuffixed words, when at least 50% of the list contained suffixed words: this ratio facilitated the use of a decompositional strategy.
- The cognitive system tries to strike an economic balance between storage space and processing energy
Full listing approaches: Lukatela’s satellite model (1987)
• Nominative nouns fastest, other cases do not differ from each other.
• Suffixed forms (both inflectional and derived) are attached to the stem like satellites

Full listing approaches
• Network models (pl. Bybee, 1995)
  – Whole words only
  – Not list metaphor \(\rightarrow\) representations are part of an associative network with mutual
    connections between phonological and semantic levels (only form and meaning)
  – No separate morphology: morphological generalisations emerge from form-meaning
    Relations \(\rightarrow\) epiphenomenon, does not have a separate level of its own
  – Generalisations are gradual and schematic \(\rightarrow\) no discrete symbolic representation

Taft: Affix-stripping model
• 1. a nonexisting combination of a real stem+a real suffix (*dejoice) >
  2. real prefix+nonreal stem (*dejouse)
• Pseudoprefixed words like relish > RT whole words like manage and argue or prefixed
  words rethink and decompose
• 1. access representation is already analysed into morphemes which reach the central
  lexicon where both morphemes are activated, it is only their combination that doesn’t
  work
• 2. We get an error signal during access already, there is nothing to activate.

Morphological processing
• Experimental evidence (Taft & Forster 1975):
  • nonword interference:
    - Real (bound) stems take longer to reject in lexical decision, than pseudostems:
      -vive (revive) vs. –lish (relish)
    - The same happens with nonwords that are a nonexistent combination of an existing
      prefix and an existing stem (dejoice) \_takes longer to reject than nonwords containing
      pseudostems (dejouse)
    - \(\rightarrow\) bound stems have their own access representations (morphological structure):
      prefixed words are decomposed during lexical access
• Morphological effects can be separated from semantic and phonological effects:
  - Tables-table: similar in both meaning and form
Multimorphemic words and holistic storage?

- For frequent words
- For non-transparent compounds (butterfly, deadline) or derived words (casualty, creature, etc.).
- Derivational suffixes are more closely associated to the stem than inflectional suffixes
  - Derivational suffixes are rarely detached from the word in speech errors
  - In lexical decision tasks, derived words do not take longer to recognize than
    pseudoderived ones (plastic, badness, shipment vs. panic, harness, garment).
  - Yet derivational suffixes can be productive

Dual route models

- AAM (Augmented Addressed Morphology)
  - Caramazza et al. (1988)
  - Whole word and morphemic representation as well
  - E.g., walked activates: /walk/ /ed/ and /walked/
  - We access each familiar word through the whole word route -> decomposition only for
    new (unknown) words -> holistic access always wins
- MRM (Morphological Race Model)
  - Schreuder & Baayen (1995)
  - Whole word and morphemic representation as well
  - Pl., walked activates: /walk/ /ed/ and /walked/
  - Recognition attempts to access the word on both routes at the same time -> parallel
    dual route model
  - Properties of the stem and suffix determine which one is the winner: e.g., Frequency,
    semantic, phonological transparency, affixhomonymy etc.

Morphologically complex words

- Analysis vs holistic storage
  - Walk-ed, smart-est, animal-s, develop-ment
  - fel-kidt-ás-nak ‘for a crying’
  - Verbal prefix-stem-derivational suffix-inflectional suffix
- Holistic storage
  - Fast access, but big storage capacity requirements
  - Productivity, i.e. recognition and formation of new forms is based on analogy, similarity
- Analytic or decompositional representation
  - A separate mechanism for analysing complex forms, only the stem needs to be stored
  - Small storage capacity requirements, but new, separate mechanism is evolutionarily
    costly
  - Productivity: rule system
- Holistic: processing time is not longer for multimorphemic words (of the same length)
  than monomorphemic words
- Analytic: there is an extra cost for analysis/composition cost besides accessing the stem;
  morphologically complex words take longer to process; more morphemes ➔ longer RTs
- The mental lexicon is an associative memory network ➔ memory effects:
  - frequency (frequent words are easier to access and process)
  - similarity (similar words behave in similar ways)
- Holistic: word frequency of multimorphemic words affects processing.
  (processing of walked is primarily affected by frequency of the word walked and not of
  the stem walk)
• Analytic: stem frequency is more relevant, as only the stem has a distinct representation; morphologically complex words are accessed through the stem (on this view, processing speed of the word walked is primarily determined by the stem frequency of walk, i.e. the sum frequency of all the words that have walk as their stem).

• In most recent models, storage is not unitary, it is affected by several different factors, which together determine whether a word is represented holistically or analytically. Mixed models are basically analytic models, in which complex words are, under certain circumstances, stored holistically.

Transparency
• Marslen-Wilson et al., 1994: Cross-modal priming
• Mutual inhibition between derivational suffixes, but not for prefixes
• Multimorphemic words are stored analytically at the level of the central, abstract level of lexical representation, and they are accessed through the stem
• Only true is the semantic relationship btw derived word and stem is transparent (sincere-sincerity vs casual-casualty)

SAID
• Based on Finnish (morphologically rich)
• Stem Allomorph Inflectional Decomposition (Niemi et al., 1994a, b; Laine et al., 1995)
• Storage is influenced by regularity and type of suffixation
  – Regulars analytically, irregulars holistically
  – Inflection analytically, derivation holistically
(Finnish patients with aphasia make more errors with inflected than with derived words).
• Morphologically complex words have both holistic and analytic representations

Cross-modal priming task
• Subjects heard spoken prime words and respond with a lexical decision task to visually presented probes flashed onto a computer screen
  – The lexical decision task was simply deciding whether or not the letter sequence flashed on the screen was a real word or a non-word. (In approximately half of the experimental trials the probe letter sequence was a phonotactically legal non-word letter sequence, e.g.: 'glark').
  – It was the morphological relation between the prime and the probe word, not any homological similarity between prime and probe that was found to matter (Marslen-Wilson, Tyler, Waksler and Older, 1994)
• When subjects heard the word 'friendly' spoken at the same time as they made a lexical decision to the visual probe: friend, their responses were faster by 40 – 60 milliseconds than to an unrelated probe, but when 'tinset' was the auditory prime for the probe word tin, no priming effect was obtained.
Test conditions and morphological priming effects

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Example: prime/probe</th>
<th>Priming effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. [+Morph. + Phon]</td>
<td>friendly/friend</td>
<td>Yes</td>
</tr>
<tr>
<td>2. [+Morph. - Phon]</td>
<td>elusive/elude</td>
<td>Yes</td>
</tr>
<tr>
<td>3. [+Morph. - Phon]</td>
<td>serenity/serene</td>
<td>Yes</td>
</tr>
<tr>
<td>4. [-Morph. + Phon]</td>
<td>tinsel/tin</td>
<td>None</td>
</tr>
</tbody>
</table>

Only condition 4 where there is a phonological similarity between prime and probe but no morphological relationship failed to yield a priming effect.

What are the effects of semantic similarities amongst the prime-probe pairs in this experiment? Can they be separated from effects of morphological similarity?

Morphological and Semantic relatedness priming effects

<table>
<thead>
<tr>
<th>Condition</th>
<th>Example: prime/probe</th>
<th>Priming effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. [-Sem. + Morph.]</td>
<td>casualty - casual</td>
<td>None</td>
</tr>
<tr>
<td>2. [+Sem. + Morph.]</td>
<td>punishment - punish</td>
<td>Yes</td>
</tr>
<tr>
<td>3. [+Sem. -Morph. - Phon]</td>
<td>idea - notion</td>
<td>Yes</td>
</tr>
</tbody>
</table>

There is no evidence here for an independent effect of priming by virtue of a morphological relationship between the prime and probe stimuli. Marslen-Wilson et al. (1994) do not describe their results this way. They say that only morphologically related pairs which are ‘semantically transparent’, or perceived to be semantically related, yield priming effects.

Evidence for an independent role for morphological structure comes from the way that semantic transparency was found to interact with prefixed versus suffix-derived morphological constructions.

Why do suffixed words fail to prime each other?
- *Confession* fails to prime *confessor*. But *unfasten* primes *refasten*. Why?
- The words *confession* and *confessor* are semantically related and this semantic similarity should yield a priming effect.
- Some countervailing inhibitory effect must be at work.
- This can only come from competition between the two derivational suffixes *-ion* and *-or*.
- We need also to explain why prefixes do not mutually inhibit one another.
- A prefix typically allows numerous stem attachments.

- ment
  - govern
  - or
### 7.2. Word-level factors influencing processing

- **phonological transparency:** controversial results, modality- and task dependent
- **semantic transparency**
  - Marslen-Wilson et al. (1994): itermodal priming -> priming for *happi+nness happy*, but not for *apart+ment-apart* -> modality independent, morphologically structured level of representation for semantically transparent words. (direct access from speech without intermediate access steps)
  - Baayen, Burani & Schreuder (1997): even semantically transparent words can be stored holistically -> semantic transparency is a necessary but not sufficient condition for morpheme-based access
  - Forster & Azuma (2000): significant masked priming for pairs like *submit-remit* -> semantic transparency might not be necessary for decomposition
  - Rastle et al. (2003): masked priming effects for *apart+ment-apart*, words made up of "pseudo-morphemes", *broth+er-broth*, but not for pairs like *brothe* → *broth* -> morphologically induced "automatic" analysis at the level of mediating access: gradual effects (at least for visual words).

#### MRM

<table>
<thead>
<tr>
<th>Condition</th>
<th>Morphological type</th>
<th>Example</th>
<th>Priming effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suffixes:</td>
<td>prime - probe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. [-Sem + Morph.]</td>
<td>derived - stem</td>
<td>successfully - succeed</td>
<td>no</td>
</tr>
<tr>
<td>2. [+Sem + Morph.]</td>
<td>derived - stem</td>
<td>punishable - punish</td>
<td>yes</td>
</tr>
<tr>
<td>3. [-Sem + Morph.]</td>
<td>derived - derived</td>
<td>controversial - controversial</td>
<td>no</td>
</tr>
<tr>
<td>4. [+Sem + Morph.]</td>
<td>derived - derived</td>
<td>cooperate - cooperate</td>
<td>no</td>
</tr>
<tr>
<td>Prefixes:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. [-Sem + Morph.]</td>
<td>derived - stem</td>
<td>remain - remain</td>
<td>yes</td>
</tr>
<tr>
<td>6. [+Sem + Morph.]</td>
<td>derived - stem</td>
<td>uncertain - ensure</td>
<td>yes</td>
</tr>
<tr>
<td>7. [-Sem + Morph.]</td>
<td>derived - derived</td>
<td>depress - express</td>
<td>no</td>
</tr>
<tr>
<td>8. [+Sem + Morph.]</td>
<td>derived - derived</td>
<td>unfavour - unfavour</td>
<td>yes</td>
</tr>
</tbody>
</table>

[Diagram of MRM]

**LEVEL**

- **SYNTACTIC/SEMANTIC**
  - syntactic representations
  - semantic representations
  - Licensing and composition

---

**LEMMA (concept nodes)**

**LEXEME**

- access representations
  - intermediate access representations
  - segmentation

---

Psycholinguistik – Folien, WS11 Seite 94
• Three-level interactive activation frame (whole-word access) with a mechanism that can perform symbolic computations on representations available at different levels (decomposition access)
• Activation feedback mechanism
• cca. Any factor that complicates morpheme-based access increases activation level of the holistic form and supports full-form access

Word-level factors influencing processing
• Frequency
  – Most robust and well-documented effect
  – Taft (1979): cumulative stem frequency influences processing of inflected verbs: the sum frequency of word+inflected form has an effect
  – Sereno & Jongman (1997): surface frequency (the frequency of the word form) affects processing of inflected nouns (in English). \(-\)\(\rightarrow\) full-form access for inflected words as well
  – Baayen et al. (1997): cumulative stem frequency (sg. + pl.) influences processing of monomorphemic (singular) nouns in Dutch
  – Size of morphological family: the number of affixed and complex word in which a word is a part
    • Schreuder & Baayen (1997) for monomorphemic words
    • Bertram et al. (2000) for inflected words
• Affix properties of the affixed word
  – Affix saliency: the probability that an affix is a processing unit in the given language
• Affix length and affix frequency: long and frequent affixes are more likely to be used in analysis than shorter and less frequent ones (overview Laudanna & Burani, 1995)
• Affix productivity: more productive affixes are more likely units of processing
• Validity: the ratio of the given string being a real vs a pseudosuffix (Italian -\(\text{\textendash}\)ri 81% -co 18%)
• Affix homonymy: Bertram et al. (2000c) Dutch –er \(~\), agentive marker ‘werker’ [worker] and intermediate ‘mooier’ [nicer], leads to holistic processing

7.3. Words and rules
• Basic dual pattern in language:
  – Rule-based system
    • mental grammar
    • procedural memory
  – Item-based system
    • lexicon
    • declarative memory
• declarative/procedural model
• Mapping from the sound structure of a word to its meaning may be achieved compositionally (by rule),
  • or by directly matching a word-form in the mental lexicon (the lexical route).
• Dual model (Pinker, 1991, 1999; Pinker–Prince, 1994: language is implemented by two systems
• grammar: categorical and rule-based
  – Computational rule system, procedural knowledge for the assembly of items, which applies to a clearly distinct and independent (syntactic or semantic) category of items
• lexicon: associative and statistical mecanisms and storage
  – The mental lexicon is a structured store of linguistic items and their idiosyncratic properties
Morphological regularity and frequency of word usage.
• Low frequency words which take regular inflections.
• Three types of inflection may be distinguished.

<table>
<thead>
<tr>
<th>Basic or non-tensed form</th>
<th>Past tense form</th>
<th>Occurrence in speech</th>
<th>Morphological type</th>
</tr>
</thead>
<tbody>
<tr>
<td>go</td>
<td>went</td>
<td>very high frequency</td>
<td>suppletive</td>
</tr>
<tr>
<td>leave</td>
<td>left</td>
<td>mid range frequency</td>
<td>partial regularity</td>
</tr>
<tr>
<td>depart</td>
<td>departed</td>
<td>low frequency</td>
<td>fully regular</td>
</tr>
</tbody>
</table>

What is the rationale for the regularity – frequency trade-off?

Rationale for the regularity – frequency trade-off?
• Low token frequency forms would take too long to retrieve
• Irregularly inflected forms are phonologically economical (shorter).
• Partially regular forms form families of phonologically related members.
  e.g.: keep [/ki:p/ - /kept/], leap, creep, sleep
• Regularly inflected forms show no family resemblances in terms of phonological similarity.

Pinker’s dual route model (1991)
• What is a rule of grammar?
  – A data structure or computational procedure implemented on neural hardware
  – An epiphenomenon of completely different neurocomputational processes
• Testing ground: inflectional morphology, mostly English past tense
  – regular: walk-walked
  – irregular: go-went, sing-sang
  – regular: fully rule-driven, typical case of a grammatical rule implemented in the brain,
    classical symbol-manipulating rule that attaches an affix to the variable representing the stem
  – Irregulars: different degrees of predictability*memory memorized word pairs
    (associations stored in memory, which show the characteristics of memory representations-frequency and similarity effects)

Traditional grammars
• Rules for regulars, rote for irregulars

  Problems:
  1. Base and marked forms are similar for irregulars - swing-swung
  2. Stems with irregular forms tend to come in clusters (they are similar to each other):
      sing, ring, spring, sink, stink
  3. Semiproduticty

- Lexical phonology
  every regularity, whatever the scope, is a rule, to minimize memory load ⇒ explains 1, but not 2 and 3

- Connectionist models
  Every process is a function of memory ⇒ cannot explain 1
Pinker's dual route model (1991)

<table>
<thead>
<tr>
<th>Adattitus</th>
<th>Rule system</th>
<th>Associative memory system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word frequency effect in recognition</td>
<td>learn affected by sum frequency of learn, learned,</td>
<td>went only affected by frequency of went, go does</td>
</tr>
<tr>
<td>Production</td>
<td>0 frequency effect</td>
<td>not count go → went is frequency dependent</td>
</tr>
<tr>
<td>Priming</td>
<td>learned → learn facilitation</td>
<td>went → go no facilitation</td>
</tr>
<tr>
<td>Nonword suffixation</td>
<td>Rule-based by default; proper names, loan words here</td>
<td>Novel words similar to irregulars: irregular pattern</td>
</tr>
<tr>
<td>Idioms</td>
<td>Regular: can change</td>
<td>Irregular: doesn’t change</td>
</tr>
<tr>
<td>Aphasia</td>
<td>Difficult (frontal, nonfluent)</td>
<td>“easier”: random</td>
</tr>
<tr>
<td>SLI</td>
<td>difficult</td>
<td>easier”: random</td>
</tr>
<tr>
<td>Williams syndrome</td>
<td>Rule easier</td>
<td>More problems with irregulars</td>
</tr>
</tbody>
</table>

Schema instead of rules

- No distinct subsystems within language, no rules, only regularities
- No real differences between background mechanisms of regulars and irregulars; apparent difference is due to frequency differences-type and token (Bybee, 2001).
- Function of rule is replaced by schema as a generalization-representation:
  - An organizational principle within the lexicon; not distinct from the word forms in its scope; a schema equals the word forms that are part of it, which can have semantic, phonological or morphological etc. connections. As a consequence the lexicon is not a random list of words with their idiosyncratic properties. Storage is structured by regularities and similarities.
- Productivity of a morphological pattern is greatly influenced by its type frequency (different for regulars and irregulars): the more item a pattern contains the more productive it is
- Productivity is gradual, not all-or-nothing.
- Type frequency in interaction with similarity bw members of the class. Gang effect: – a sing-sang irregular type, with greatly similar members (sing, ring, string, fling) is extended to nonwords that are similar to members of the gang (spling-splang) (Bybee–Moder, 1983).
- Irregular types show only limited productivity, and extent of productivity is determined by the same parameters in dual models as well.
- Experimental evidence that the behavior of regulars is affected by their frequency.
- Stemberger és MacWhinney (1986) error elicitation in production of past tense forms of regular verbs. Regular forms are also stored in the lexicon, and frequency affects the strength of the representation
- Alegre and Gordon (1999) found similar frequency effects for regulars in a lexical decision task
- Albright and Hayes (2003) same for well-formedness judgments on nonwords
Ullman: The procedural/declarative model

- **procedural memory:**
  - Learning and control of motor and cognitive skills and habits
  - Rule learning
    - Especially important in learning sequences
  - Slow learning – fast and automatic retrieval
  - Implicit: not available to consciousness

**Other functions**

- grammar
- lexical access
- Dynamic mental imagery
- Working memory
- Rapid temporal processing

- These are not all part of procedural memory, but can be based on the structures on which it is based too.

**Brain structures**

- Cortex: mainly in left hemisphere
- Frontal-basal ganglia are important – connected intercortical structures:
  - Neostriatum
    - Nucleus caudatus
    - Putamen
  - Globus pallidus
  - Nucleus subthalamicus
  - Substantia nigra

8. **SENTENCE PROCESSING**

**Complications in sentence processing**

- How should we incorporate ambiguous words into a sentence?
- Do we keep various options open and postpone decision making to later?
- Do we commit to one interpretation early on, based on parsing principles, and only later fix mistakes?
- No rule allows two adjacent NPs at the beginning of the sentence. The sentence must be reanalyzed.

```
NP \rightarrow \text{det (adj)} \text{ N}
S \rightarrow \text{NP VP}
old \rightarrow \text{A}
old \rightarrow \text{N}
man \rightarrow \text{N}
man \rightarrow \text{V}
```
8.1. Parsing

- The syntactic analyser or “parser”
  - Main task: To construct a syntactic structure from the words of the sentence as they arrive
  - Main research question: how does the parser “make decisions” about what structure to build?

Sentence Comprehension
- Modular

**Different approaches**

- **Immediacy Principle:**
  access the meaning/syntax of the word and fit it into the syntactic structure
  - **Serial Analysis (Modular):** Build just one based on syntactic information and continue to try to add to it as long as this is still possible
  - **Interactive Analysis:** Use multiple levels (both syntax and semantics) of information to build the “best” structure

- **Parallel Analysis:** Build both alternative structures at the same time
- **Minimal Commitment:** Stop building - and wait until later material clarifies which analysis is the correct one.

Sentence Comprehension
- A vast amount of research focuses on: **Garden path sentences**
  - A garden path sentence invites the listener to consider one possible parse, and then at the end forces him to abandon this parse in favor of another.
- Garden path sentences
  - The horse (1) raced (2) past (3) the barn (4) fell (5).
Questions
- Is structure-building immediate, or does the parser wait for later and more information?
- Are top-down effects allowed, or is the parser strictly bottom-up?
- Top-down effects before hearing the acoustic signal, or only after activating the list of potential words?
- Computers only one syntactic and/or semantic analysis or works with several possible interpretations in parallel?

The autonomy of syntax (Chomsky)
- We process language serially
- phonology → lexical processing → syntactic processing → semantic processing
- semantic analysis starts only when we reach a clause boundary

8.2. The Derivational Theory of Complexity
- Testing Chomsky’s Theory
- Transformationally complex sentences are more difficult to process
- Miller and McKean (1964): do a transformation, than find the resulting sentence in a list
  - The robot shoots the ghost. 0
  - The ghost is hot by the robot. 1
  - Is the ghost not shot by the robot 3
- Linear relationship between number of transformations and time
- Support from other experiments as well: participants remember fewer words in a dual task when the sentences are more complex → requires more memory resources
Miller (1962)

- 1. Mary hit Mark. K(ernel)
- 2. Mary did not hit Mark. N(egative)
- 3. Mark was hit by Mary. P(ositiv); (Passive)
- 4. Did Mary hit Mark? Q(uestion) (interrogative)
- 5. Mark was not hit by Mary. NP
- 6. Didn’t Mary hit Mark? NQ
- 7. Was Mark hit by Mary? PQ
- 8. Wasn’t Mark hit by Mary? PNQ

Transformational cube

The Derivational Theory of Complexity

- Miller & McKean (1964): Pairing of sentences with the same meaning or ‘kernel’
  - Joe warned the old woman.
  - The old woman was warned by Joe. P 1.65s
  - Joe warned the old woman.
  - Joe didn’t warn the old woman. N 1.40s
  - Joe warned the old woman.
  - The old woman wasn’t warned by Joe. PN 3.12s

“The initial results were breathtaking. The amount of time it takes to produce a sentence, given another variant of it, is a function of the distance between them on the sentence cube. (Miller & McKean 1964.)

“...It is hard to convey how exciting these developments were. It appeared that there was to be a continuing direct connection between linguistic and psychological research. [...] The golden age had arrived.”

- Offline tests, explicit instructions for transformations—online analysis can be quite different
- Slobin (1966): derivational complexity effects are only obtained when there is no semantic cue (i.e. only in reversible passives):
  – The ghost was chased by the robot.
  – The water was flowered by the robot.
- Thomas Bever (2001): the attempt to use transformational grammar directly as a processing model failed

8.3. Modular approaches

- Analysis only builds on syntactic phrase structure
- Ambiguities are resolved by specific principles of syntactic analysis without invoking any other system
- These principles account for all ambiguities
- Revision: the parses often makes mistakes, and has to revise the initial analysis (with the help of thematic and discourse information).
- A two-stage analysis: an initial analysis builds only on syntactic information, the second revision takes into account other kinds of information
- serial vs parallel autonomous model: initially only one or all possible syntactic analyses
8.4. Interactive approaches

- All levels of sentence processing cooperate freely and continuously
  (E.g. Crain and Steedman’s referential theory)
- Computing several analysis based on parallel competing constraints
- Constraints can be non-syntactic as well (e.g. discourse context): what we already know about the sentence, influences comprehension
- The parser tries to assimilate the input into a discourse model
- Several analyses are computed and compared on the basis of their fit to the discourse model
- Preferences without context: minimal presuppositions that are necessary for incorporation into the discourse model
- Predicts that preferences change in context

**Constraint satisfaction model**
Several potential analyses are activated, although we are only aware of one. At the end of the sentence, when it turns out to be wrong, we switch to another (Tanenhaus, Trueswell). Both interpretations are activated (e.g. in case of a word, both the nominal and the verbal meanings), unconsciously. When the conscious analysis fails, we become aware of the other one. It does not predict that ALL possible interpretations are activated: semantic context and expectations influence alternative interpretations.

**Effects**
1. Predictability
2. Syntactic processing preferences
3. Frequency
4. Context
5. Thematic fit
6. Memory constraints

**Word Order (predictability)**
- Syntax: probabilities of various words occurring together?
- **Miller and Selfridge (1950)** created word lists that were various approximations to English.
- For a second-order approximation, present a word to a person and have them use it in a sentence. See what word they give after the target word, give that to someone, see what word they use, etc. When you string these together, you have a sequence that is a second-order approximation.
- Scale up for 3-7
- Miller and Selfridge (1950):
  - First order:
    - 30: house reins women brought screaming especially much was said cake love that school to a they in is the home think with are his before want square of the wants
  - Third order:
    - 30: happened to see Europe again is that trip to the end is coming here tomorrow after the packages arrived yesterday brought good cheer at Christmas it is raining outside as

(Music in Comedy: Monty Python/Finishing Sentences (1974)
http://www.youtube.com/watch?v=r_cRP6MhM8k&playnext=1&list=PLEF3A8DB59699CDAD)
– Fifth order:
30: go it will be pleasant to you when I am near the table in the dining room was crowded with people it crashed into were screaming that they had been
– Seventh order:
30: won’t do for the members what they most wanted in the course an interesting professor gave I went to at one o’clock stopped at his front door and rang the

– Look at recall of the lists.
– Does approximation to English affect recall?

– Order of approximation does affect memory.
– Could something like this be scaled up to account for syntax?
– Or, does understanding syntax require something more?

Predictability
Grammatical structures and regularities make it possible for us to make statistical predictions during speech processing. The 50 most frequent English words make up 60% of speech and 45% percent of written language. We utter 10-15 words before repeating a word.

What do we remember?
Surface structure in short term memory, long-term memory: only meaning
• Sachs (1967): Participants listened to paragraphs of texts, which contained a target sentence, *He sent a letter about it to Galileo, the great Italian scientist*. A beep during listening, participants had to decide whether they heard the same sentence or not.
• Evidence for gist over verbatim (Sachs, 1967):
  A. He sent a letter about it to Galileo, the great Italian scientist.
  B. He sent Galileo, the great Italian scientist, a letter about it.
  C. A letter about it was sent to Galileo, the great Italian scientist.
  D. Galileo, the great Italian scientist, sent him a letter about it.
• The greater the distance between original and test sentence, the less correct they are, the less sensitive they are to structural variations (b and c), but they still notice semantic changes
• Sometimes structure is preserved (e.g. in jokes), but we mostly only remember the gist
Early theories

• Language is processed serially, phonology, lexical processing, syntactic processing and semantic processing are strictly ordered, and syntactic processing is independent of semantic analysis (the autonomy of syntax) (Garrett, Fodor and Bever, 1966)
• Semantic analysis does not start before we reach the clause boundary

Units of analysis

• Chunks or clauses
  – *I was going to take a train to New York, but I decided it would be too heavy*
    First we store the meaning of the first one, than that of the second one, than we try to integrate the meaning of the two.
• A popular method: clicks are placed under the sentence; participants have to tell where the click was (Fodor and Bever, 1965, Garrett, Fodor and Bever, 1966)
  – *That he was happy was evident from the way he smiled*
• They were more precise when the clicks were at clause- or phrase boundaries, in other cases, clicks wandered to clause boundaries
• The same pattern, when intonation patterns were eliminated from sentences. The sentence is a perceptual unit based on its syntactic structure only
• Is prosody not important?
• Have clicks been perceived in the wrong place, or they wandered away in memory?
• The migration effect is decreased, but is not completely eliminated when the task is to mark the clicks in the text during listening instead of recall. Prosody does matter.
• Argument supporting the memory effect interpretation: the placed clicks to clause boundaries even when there were none (but participants were told that there are subliminal clicks; Reber and Anderson, 1970)
• Stine (1991), self-paced reading:
  – *The Chinese, who used to produce kites, used them in order to carry ropes across the river.*
• There is a pike after the first NP (The Chinese) (RT slows down), than RT drops and starts to rise continuously until the clause boundary (kites), than rises on until ropes.
• The more clauses a sentence contains, the higher every new pike is.
• The same pattern in the elderly, they chunk the sentence in the same way, but they do not show increasingly higher pikes: they are not as effective in the integration of information.

Fodor, Bever and Garrett (1974)

• Goal of analysis: to recover the underlying structure
• Perceptual heuristics: use of surface structural cues
• Parsing strategies
  • e.g. hearing the or a we know we are at the beginning of an NP
• Strategy of canonical sentences: NVN often maps onto SVO structure.
• Builds only on syntactic cues

Kimball (1973)

• Surface cues invoke strategies
• 7 principles, on which initial computation of structure is built
• These principles are driven by psychological constraints like minimizing memory load
8.5. Kimball’s 7 principles

1. Parsing in natural language proceeds according to a top-down algorithm:
   – We start from the sentence node and wait for words that correspond to the given phrase type (e.g. when we see a determiner, we expect to see a noun in a short time)

2. Right association:
   – Terminals associate to the lowest non-terminal node i.e. to the constituent under construction
   – Sentences organize into right-branching structures (perceptually less complex than leftbranching or center-embedded)
     • The policeman saw the thief with the telescope.
     – Right association: telescope is thief’s rather than policeman’s

3. New nodes (Knotenpunkt, Knoten):
   – the construction of a new node is signalled by a function word. Deleting complementizers and relative pronouns make sentences more complex
     • A. He knew the girl left.
     • B. He knew that the girl left.
     • a. A kisfiú megkérdezte az anyját merre találja.
     • b. A kisfiú megkérdezte hogy az anyját merre találja.

4. Two sentences:
   – The processor can handle maximum 2 sentences at a time. That is why central embedding is difficult:
     • a. The vampire the ghost the witch liked loved died.
     • b. The witch liked the ghost that loved the vampire that died.

5. Closure:
   – A phrase is closed as soon as possible (unless the next node can be analysed as a constituent of the phrase)

6. Fixed structure:
   – It is costly to reorganize the constituents after a phrase has been closed.
   – Garden path sentences are difficult.
     • After the child had visited the doctor prescribed a course of injections

7. Processing:
   – When a phrase is closed, it is pushed down into a syntactic processing stage and cleared from STM (short-term memory). Structure is quickly forgotten.

8.6. Ambiguity

• Local:
  – Since Jay always jogs a mile seems like a short distance to him.
  – After the child had visited the doctor prescribed a course of injections.
  – The witness examined by the lawyer turned out to be unreliable.

• Persistent:
  – The policeman saw the thief with the binoculars.
  – Visiting relatives can be boring.

• Source of ambiguity:

• Sources:
  1. Lexical ambiguity in meaning and/or in syntactic category (The witness examined…).
  2. Ambiguous dependency (The policeman saw the thief with the binoculars).
Structural ambiguity

Garden path phenomena
Suggest that the parser commits to one structure (which has to be revised in the face of contradicting evidence later in the sentence)
*The horse raced past the barn fell.*
*The student forgot the answer was in the book.*

Right association

Preferred analysis: left (? For German? Hungarian?)
Kimball (1973) **New nodes:** the construction of a new node is signalled by a function word

Why central embedding is difficult

a. The girl the man kissed left.
b. The man the boy saw kissed the girl.
c. # The girl the man the boy saw kissed left.

Violates the principles of Right association, New nodes and Two sentences

8.7. **Frazier and Fodor (1978)**

- The sausage machine
- Makes Kimball’s model simpler, the principle of processing
- PPP: preliminary phrase packager: only syntactic cues, a six-word window, simpler structure, Minimal Attachment
- SSS: sentence structure supervisor
- Wanner (1980) problematic six-word sentences:
  – Vampires werewolves rats kiss love sleep.
  – Vlad said that Boris cried yesterday.
- Fodor and Frazier (1980) added the principle of right attachment: the Garden Path Model
  All Kimball-principles summarized in 2 in the Garden Path model
- Minimal attachment: Prefer the interpretation that is accompanied by the **simplest**
  structure.
- simplest = fewest branchings (tree metaphor!)
- Count the number of nodes = branching points
- Late closure: Incorporate incoming material into the phrase or clause currently being processed.
'Garden Path' model (e.g. Frazier & Rayner, 1982, 1987)

A special case of Minimal attachment.

Two-stage serial model

- Build trees using syntactic cues:
  - phrase structure rules (universal grammar)
  - plus two parsing principles
    - Minimal Attachment
    - Late Closure

- Go back and revise the syntax if later semantic information suggests things were wrong
  - Since Jay always jogs a mile seems like a short distance to him.
  - After the child had visited the doctor prescribed a course of injections.
  - The evidence examined by the lawyer turned out to be unreliable

A serial model

- Minimal Attachment
  - Prefer the interpretation that is accompanied by the simplest structure.
    - simplest = fewest branchings (tree metaphor!)
    - Count the number of nodes = branching points
      The girl hit the man with the umbrella.

- Late Closure
  - Incorporate incoming material into the phrase or clause currently being processed.
    OR
  - Associate incoming material with the most recent material possible.
    She said he tickled her yesterday
Late closure

Since Jay always jogs a mile seems like a short distance to him.
Phrase is closed after mile, rather than after jogs

Closed late:
Since Jay always jogs a mile, this seems like a short distance to him.
Closed early:
Since Jay always jogs, a mile seems like a short distance to him.

→ sentences that are closed late are easier to understand than those closed early. In sentences with early closure, the processing time of the disambiguating word increases.
Empirical results: minimal attachment
Frazier and Rayner (1982), exe-tracking:
Readers face a processing difficulty as soon as they encounter evidence in favor of the nonminimally attached alternative
Compare:
a. The pupil spotted by the proctor was expelled - ambiguous
b. The pupil who was spotted by the proctor was expelled - unambiguous
c. The pupil seen by the proctor was expelled - unambiguous
At by the proctor fixation times increase in ambiguous sentences relative to unambiguous ones.
In reading, regressive eye movements are more likely to occur in these cases.
Sentences disambiguated according to the principle of minimal attachment do not show such differences relative to controls (The pupil spotted the proctor).
• Similar results to other kinds of structural ambiguity

Minimal attachment
Garden path sentences (Rayner & Frazier, ‘83)

The spy saw the cop with a telescope.

The spy saw the cop with a revolver.
Interactive Models
Other factors (e.g., semantic context, co-occurrence of usage & expectation) may provide cues about the likely interpretation of a sentence
Trueswell et al (1994):
- The evidence examined by the lawyer...
- The defendant examined by the lawyer...

...evidence typically gets examined, but can’t do the examining

Garden Path model
• Emphasizes the operation of universal principles → compatible with the idea of an innate universal grammar. No reason to suppose that universal processing principles have to be relearned by each individual.
• Clear and simple universal principles (economy)
• Principles are testable and falsifiable in most languages
• But results are based on studying only a few structures
• Focuses too much on ambiguous structures, which is not necessarily the same as language processing

Interactive Models
• Other factors (e.g., semantic context, co-occurrence of usage & expectation) may provide cues about the likely interpretation of a sentence
Trueswell et al (1994):
- The evidence examined by the lawyer …
- The defendant examined by the lawyer…

... A defendant can be examined but can also do examining.

**Semantic expectations**
Other factors (e.g., semantic context, co-occurrence of usage & expectation) may provide cues about the likely interpretation of a sentence.

Taraban & McCelland (1988)
– Expectation
The couple admired the house with a friend but knew that it was over-priced.
The couple admired the house with a garden but knew that it was over-priced.

---

8.8. **What about spoken sentences?**
All of the previous research focused on reading, what about parsing of speech?
– Methodological limits – ear analogue of eye-movements not well developed
  • Auditory moving window
  • Reading while listening
  • Looking at a scene while listening
– Some research on use of intonation

**Chunking, or “phrasing”**
A1: I met Mary and Elena’s mother at the mall yesterday.
A2: I met Mary and Elena’s mother at the mall yesterday.
Phrasing can disambiguate

One intonation phrase with relatively flat overall pitch range.

Separate phrases, with expanded pitch movements.

Summing up
- Is ambiguity resolution a problem in real life?
  Yes (Try to think of a sentence that isn’t partially ambiguous)
- Many factors might influence the process of making sense of a string of words.
  (e.g. syntax, semantics, context, intonation, co-occurrence of words, frequency of usage, ..)

Garden path effects in fact semantic effects?
- Milne (1982): a sentence is only GP when it generates semantic expectations that are violated later. Only a. is GP:
  a. The granite rocks during the earthquake.
  b. The granite rocks were by the seashore.
  c. The table rocks during the earthquake.

- Crain and Steedman (1985) there is no such thing as a semantically neutral context, world knowledge always has an influence:
  – Rapid grammaticality judgments
  – Starting from the general observation that GP sentences are judged (erroneously)
    ungrammatical significantly more frequently
  – Semantically less plausible sentences (a.) – supporting semantic effects:
    a. The teachers taught by the Berlitz method passed the test.
    b. The children taught by the Berlitz method passed the test.
Crain and Steedman’s Referential Theory
• Parsing preferences are not explained by syntactic constraints, but by semantic effects: The more presuppositions you need, the more difficult processing becomes.
• Syntactic representation is built incrementally: at every new word several analysis are generated in parallel, and contexts selects from among them.
• Discourse context disambiguates immediately—this is an important difference between Referential Theory and Garden Path theory.
• In this theory it is only referential complexity in the discourse model that determines processing difficulty. In constraint-based theories all kinds of semantic effects count, including world-knowledge.

1) The horse raced past the barn fell.
2) The horse raced past the barn quickly.
→ 2 rather than 1 not because of simpler structure, but important referential differences..
→ 1. complex NP, modified by a reduced relative clause
→ 2. Simple NP, subject
→ 1. Complex NP-k presupposes the existence of a richer set of entities, from which the relative clause selects one
→ 2. Presupposes only one entity, which has the property of being a horse.

Discourse context
Sentences never come out of nothing, the reader or listener builds a mental model or discourse model: events, entities, etc.
Ambiguity resolution has different discourse gains
a. Main Clause: introduces a new event into the model
   The lookout spotted the sniper.
b. reduced relative: a given event, selects an entity from the group
   The lookout spotted by the sniper was chastised.
Context influences ambiguity resolution:
All of our lookouts failed at keeping out of sight; one was seen by a spy, one by an enemy aircraft, and one by a sniper. The lookout spotted by the sniper was chastised.

Semantics: animacy and thematic fit
Semantics does make a difference
– The girl hit the man with the umbrella.
– The girl saw the man with the umbrella.
– The girl hit the man with the bowtie.
– The witness examined by the lawyer was useless
– The evidence examined by the lawyer was useless

Inanimates are bad Agents, and bias (tendieren) towards the reduced relative reading. The sentence beginning with The evidence... -- shorter processing time.

Frequency
• Like lexical access, frequency also has an important role in syntactic processing
• Contingent Frequency: the frequency of a given word/morpheme in a given context
  – E.g. how often spotted is an active vs passive past tense verb following a definite NP-t (the lookout spotted)
**Conflict: Frequency vs. attachment**

- Reading difficulty in the underlined positions:
  
  *Since Jay always jogs a mile seems like a short distance to him.*  
  *After the child had visited the doctor prescribed a course of injections.*

- Mitchell (1987): the difficulty also occurs with intransitive verbs, but processing slows down at the object itself:
  
  – a. *After the child had visited the doctor prescribed a course of injections.*

- Attachment based on category, followed by lexical filtering

- Frequency distributions show that the NP following the verb is generally the object of the verb. We can expect frequency effects based on verbs as a class also.

- If frequency matters, then processing difficulty should correlate with the frequency of the individual verb

- 2 factors: structural (e.g. V NP) type frequency, verb token frequency

- Predictions have been confirmed (e.g., Juliano and Tanenhaus 1993).

**Further complications**

a. *The reporter who the senator attacked admitted the error.*

b. *The reporter who attacked the senator admitted the error.*

- Subject relative clauses (b) are easier to process than object relative clauses

- The difference is not caused by plausibility or a lexical bias (both are equally plausible)

- No ambiguity resolution → theories of ambiguity resolution do not explain the difference

**Later results**

1. Trueswell, Tanenhaus, and Garnsey (1994): more constraining materials show animacy effects: more processing difficulty with very bad agents than with unambiguous controls

2. Syntactic alternatives also have an effect (how many arguments a verb has and what these are).

3. Frequency: thematic effects are strongest with verbs that are often used as passive participles and weakest with those that most often appear in simple past (MacDonald, Pearlmutter, and Seidenberg 1994).

4. There are parsing preferences, but context and frequency have a direct effect on analysis (as constraint-interaction theories claim)

**8.9. Short-term memory and language processing**

- Differences between good and poor readers can be STM differences

- Reading span task: more complex sentences _smaller spans_

- Central embedding is especially difficult

- Object relatives are more difficult than subject relatives

- Participants with smaller STM capacities perform more poorly in comprehension tasks

**Capacity effects**

- syntactic complexity

- ambiguities

Claim: Participants with smaller capacity have difficulties keeping both interpretations in mind, so if there is no contextual bias in the sentence towards one interpretation early on, they will be slower if the end of the sentence pushes the sentence to the nondominant meaning.
Capacity constraints

- Garden Path sentences: it is difficult to maintain two interpretations in parallel with smaller capacities → longer reading times, more interpretation errors towards the dominant meaning.
- Anaphor interpretation: bigger distance → greater difficulty.
  - *The lion escorted the zebra to the party, then it jumped into the car.*
  - *The lion escorted the zebra to fabulous party, then it jumped into the car.*
- Taking animacy into account is capacity-dependent: participants with bigger capacity take animacy into account.
- Adult-child differences: adults take discourse context into account, children do not.
- Complex sentences: integration of meaning at every clause boundary, also with the previous clause, elder people are not as successful as younger adults.

8.10. The Competition model

A functional model in which forms are evolve, change and are acquired to serve communicative functions. A theory of language use, like the Garden Path model, CM is also a precursor of several later interactive models, in which the parser uses different types of information in parallel:
- Syntactic, semantic, discourse and frequency information
- These kinds of information are constraints and these kind of models are constraint-based models.

Language users use the specific cues that are available in their language: Word-order, Morphology, Agreement, etc. (e.g. Bates & MacWhinney, 1982; Bates, MacWhinney & Kliegl, 1984).
CM: main principles

- **Two-stage direct mapping**
  a) **Functional level**: intention and meaning of the utterance
  b) **formal level**: surface structure of the expressions

- **Variety of form-meaning mappings**
  1 form = 1 meaning rare

- **form-meaning coalitions and their breakup**
  - Form-meaning pairs are not random, certain things tend to co-occur naturally, e.g. agent, actor, topic with the category of subject
  - In certain cases we need different criteria for mapping; e.g. if the object is topicalized, the subject needs a different mapping criterion (by X)

- **Competition**
  - Processing is based on the principle of parallel processing, where the stronger cue wins, but the other processes carry on

- **Cue-strength and cue-validity**
  - CM is relative, since the basic principles of processing are language-specific, each language can have a different order of relative importance
  - This is more compatible with a general learning mechanism
  - Strengths:
    - Crosslinguistic approach
    - Structural differences of languages affect processing
    - Language experience affects processing
  - Weaknesses:
    - Too strong to be empirically testable (?), since it can explain any pattern of results
      $\rightarrow$ unfalsifiable

---

**Crosslinguistic differences in processing: the Competition model**
(Bates and MacWhinney 1989)

<table>
<thead>
<tr>
<th>Language</th>
<th>Children</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungarian</td>
<td>Animacy &gt; Suffix &gt; Word-order</td>
<td>Suffix &gt; Word-order &gt; Agreement &gt; Animacy</td>
</tr>
<tr>
<td>Turkish</td>
<td>Suffix &gt; Word-order ? Animacy ?</td>
<td>Suffix &gt; Animacy &gt; Word-order</td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td>Animacy &gt; Suffix &gt; Word-order &gt; Agreement</td>
<td>Suffix &gt; Agreement &gt; Animacy &gt; Word-order</td>
</tr>
<tr>
<td>Warlpiri</td>
<td>Animacy &gt; Suffix &gt; Word-order</td>
<td>Suffix &gt; Animacy &gt; Word-order</td>
</tr>
<tr>
<td>Dutch</td>
<td>Word-order &gt; Suffix &gt; Animacy</td>
<td>Suffix &gt; Word-order &gt; Animacy</td>
</tr>
<tr>
<td>English</td>
<td>Word-order &gt; Animacy &gt; Agreement</td>
<td>Word-order &gt; Animacy &gt; Agreement</td>
</tr>
<tr>
<td>French</td>
<td>Word-order &gt; Animacy</td>
<td>Agreement &gt; Animacy &gt; Word-order</td>
</tr>
</tbody>
</table>

**Relative importance of cues in the assignment of the agent role in different language**
(Bates and MacWhinney 1989; Pléh 2003)
9. SPEECH PERCEPTION

- Segmentation and identification
- Not just auditory perception
- Species-specific
- Lateralized
- Categorical perception
- Newborn perceptual preferences
- Nonspeech sounds: 1,5/6 sequence, speech
- Research originates in 20th century communication and military industry
- Many groundbreaking results from devices developed for speech synthesis
- First device to decode and produce speech: vocoder.
- Principles used in its design helped to develop the sound spectrograph
- Analyses and maps sound signals on a graph, yielding a precise visual speech diagram called the spectrogram
- Present only briefly and once (in contrast to visual presentation)
- Segmentation boundaries are not clear, segments and even words overlap
  (the problem of segmentation)
- Properties of speech sounds vary, e.g. according to context (the problem of invariance)
- Yet we are quite good at it; if we hear speech, we cannot not understand it

9.1. Problems

The problem of lack of invariance
- Acoustic features of speech sounds change rapidly and significantly
- Individual differences between speakers
- According to context (e.g. in b but, bit, boom)
- Co-articulation: speaker pronounces sounds faster, listener maps same info to more than one segment
- Acoustic features of continuous speech are messy
- Beyond lack of invariance, under-articulation can be a problem as well:
  words may lose a significant proportion of information that identifies them.
- Lexical, syntactic and contextual info helps speech processing

The problem of segmentation
- Not only with sounds, but also with words
- Ice cream I scream
- Nitrate Night rate
- Take Gray to London Take Greater London
- Mapping of acoustic signals (that we see on a spectrogram) to phonetic representations is not simple
- We identify probable word boundaries in the speech stream (a prelexical process, possible word constraint: all units of segmentation should correspond to a meaningful word in the language).
- These places are also determined by the language’s rhythm and phonotactic structure
9.2. Categorical perception

- http://www.ling.gu.se/~anders/KatPer/Applet/test.eng.html
  ©Anders Eriksson
- Perception of the same difference is difficult with the category but easy across categories
- We perceive speech sounds categorically (or at least consonants)
- We perceive nonspeech sounds noncategorically (?)
- Babies show evidence of categorical perception very early after birth (Eimas, Siqueland, Jusczyk and Vigorito, 1971)
- Categorical perception studies show that there is a top-down effect in phoneme identification

Synthesized stop sounds—continuous change in place of articulation dimension
Percepts: /b/ vs /d/ vs /g/:

Liberman, Harris, Hoffman, and Griffith (1957)
*Journal of Experimental Psychology* 54, 358-368

Categorical perception: /r/ versus /l/
- http://www.wadsworthmedia.com/psychology/now/perception/labs/cate_per.swf

Categorical Perception in Speech

The phonetic boundary divides the auditory percept into two categories: /rake/ and /lake/. Did the demo work?... It should look like the sample data below:
Categorical perception of voicing in stops
• Majority of voiced-voiceless distinction is voice onset time (VOT)
• VOT: 30 msecs from air release in voiced, between 40 and 100 msecs in voiceless consonants.
• Synthesized sounds, only difference is VOT, from -40 ms to +100 ms in 10 ms steps

Species specificity?
• Monkey auditory cortex: cells respond to features that are important in human speech perception
  – (1) time delay between start of sound and the vibration of the vocal cords. (VOT)
  – (2) acoustic context of the sound (consonant-vowel co-articulation cues)
  – (3) rate of frequency changes
• Chinchillas too are sensitive to differences in voicing and place of articulation

Categorical perception
• Eimas and Corbit (1973)
• Synthesized syllables, along the voicing continuum
• ‘b’ for several minutes: sounds around the boundary are perceived as ‘p’ a lot more
• Tiring in ‘voiced’ detectors
• Selective adaptation
• It is possible that feature detectors exist, but they do not explain speech perception

Development of speech perception
What do babies like?
– They prefer speech to all other kinds of acoustic stimuli. Biologically wired to process speech.
– Motherese: slower, higher pitch, intonation.
– DeCasper and Fifer (1980) 3 day-old babies suck more (sensory apparatus to detect sucking rate changes) to hear their mother’s voice
– DeCasper and Spence (1986) The cat in the hat study: mothers read the tale aloud for 6 weeks before birth, 2-day-olds change their sucking rate to hear the story
• Mehler et al. (1988) distinguish mother tongue from foreign language
• They remember something of the tempo and rhythm → they hear speech before birth.
• This is not enough for speech yet. They are also able to distinguish phonemes after birth

Innate ability to discriminate speech sounds?
• Present very early.
• Regardless of the specific language, distinctions are perceived at the same point of the continuum, even when this distinction is not exploited by the language
• E.g. Spanish babies discriminate between p and b, although it is not contrastive
• Phoneme inventories become limited to phonemes of the mother tongue between 6 and 12 months. Gradual elimination, most can be brought back for quite some time with learning
• Categorical perception can be brought back for a long time, but the ability to perceive the lot of different ‘pa’-s as the same is also important
Development of speech perception
- Eimas et al. 1971: 1 month-old babies ba-pa 123456, dishabituation study, 123 ba, 456 pa, they did not dishabituate to 2 and 3, but they did to any of 456
- → categorical perception to speech sounds (only evidence for voiced-voiceless contrast)
- Earlier: the ability to discriminate speech sounds is the result of perceptual learning, but this evidence suggested that it is innate
- The ability to discriminate speech sounds that are not part of the native phoneme inventory disappears between 6 and 12 months
- Babies are sensitive to potential switching points, and only loose them later (like Thai babies the 30 msec boundary in VOT)
- Babies are initially sensitive to typical switching points initially (English) and they develop those that are atypical later (e.g. Thai).
- Our perceptual experience determines which perceptual categories are created
- Two different developmental processes: selection vs construction.
  Experimental evidence supports the latter
  - Lasky et al. 1975: babies discriminate English vs. Thai VOT, but not Spanish
  - Streeter ~1976: Kikuyu babies perceive English, but not adult Kikuyu boundaries.
- A combination of Universal Theory, Tuning Theory and Perceptual Learning Theory is the best model (Aslin és Pisoni, 1980)

Categorical perception
- It is not true that we are not sensitive to within-category differences: Pisoni and Tash, 1974: participants decide faster whether two ba-s are identical, when they are actually identical, not just from within the same category
- Theories of categorical perception are not as dominant as they used to be, greater role to continuous perception
- Evidence for categorical sensation is a lot weaker

Context effects
- Speech perception bottom-up or top-down?
- If word or sentence context has an effect → interactive
- Phoneme identification is easier in a word or sentence that in isolation
- Word boundaries are illusory, no correspondent in the acoustic signal

The effect of visual info on speech perception (The McGurk effect)
- Visual:: GA, Auditory: BA
- Percept: DA (McGurk & MacDonald 1976)
- Visual cues affect how we perceive certain sounds. Multimodal speech perception
- Todd (1977): tough+hole=towel
- Green et al (1991): works even with speakers of different gender
Context effect: categorical perception
• Context shifts category boundaries
• Ganong (1980) g-k, inserted in front of _iss
• Earlier g is perceived as k
• Lexical identification effect
• Connine and Clifton (1987) the effect disappears at either end of the continuum
• Connine (1990) lexical effect is perceptual, but the effect of sentence context is postperceptual (it is affected by rewarding correct answers and punishing incorrect ones)

Context and expectations
• Speech rate: a certain syllable might be perceived as /ba/ when they are embedded in a sentence with slow rate, but as /pa/ in a sentence with a fast rate (Summerfield, 1975).

Ladefoged and Broadbent (1957)
• “Please say what this word is: bit bet bat but”
Same target word only the formant frequency of the carrier sentence is different

F1 of the sentence
200-380 Hz
380-660 Hz

• Speech is robust, full of redundancies; we understand is fairly well in noise
• Comprehension is easier when it is predictable. The more predictable it is, the more background noise is tolerated

Phoneme restoration
• [http://www4.uwm.edu/APL/demonstrations.html](http://www4.uwm.edu/APL/demonstrations.html)

Context effect: phoneme restoration
• A speech sound is replaced with noise (cough or buzz), the listener thinks to have actually heard the sound (Warren, 1970). They are unable to locate the noise.
• The state governors met with their respective legislatures convening in the capital city.
• Even legislature works
• Type of masking sound does not matter, but silence is not restored
• Even more dramatic effect, with varying only the context:
  – It was found that the *eel was on the shoe.
  – It was found that the *eel was on the table.
  – It was found that the *eel was on the orange.
  – It was found that the *eel was on the axle.
• Do we actually perceive the phoneme, or do we only make post hoc guesses (unconsciously)?
• If it is actually a perceptual effect, there is no difference in percept, listeners hear phoneme+noise
• If it is post perceptual, the two are going to be different
• → Samuel: lexical context perceptual, sentence context post perceptual

– The travelers found horrible bats in cavern/tavern when they visited
– The travelers found horrible food in cavern/tavern when they visited
• First should yield better restoration for cavern, but it is not what happens, sentence context only at a later processing stage
• Samuel: selective adaptation works with restored phonemes as well:
  alpha*et 40x, than bi-di identification
• There is a top-down effect, but a restricted one, sentence context small effect

Grammarality effects
Knowledge of grammar guides our perception of speech.
White Noise: a sound stimulus that contains all spatial frequencies, sounds like your detuned television set.

Miller and Isard shadowing experiment (1963):
- grammatical sentences + noise: 63%
  Gadgets simplify work around the house..
- grammatical, nonsensical sentences + noise: 22%
  Gadgets kill passengers from the eyes..
- ungrammatical word strings + noise: 3%
  Between gadgets highways passengers the steal.
10. LANGUAGE PRODUCTION

Speech production
- Less attention than to comprehension
- Within production, mostly speech
- For long, most evidence from speech errors (natural or induced)
- 15 sounds/seconds $\Rightarrow$ 2 - 3 words
- Automatic, we cannot tell how we do it: ‘we cannot think in the middle of a word that ‘I’m gonna say /t/ or /d/ (Levelt)’
- Speech is the most central human activity but first book only in 1989: ‘Levelt: Speaking: from Intention to articulation’

Production
- Half of our language ability:
  - Input to comprehension
  - More difficult than comprehension? Comprehension leads in acquisition.
- The task:
  - To express the unordered conceptual message through a sequence of sounds
  - Taking many constraints (e.g. grammar) into consideration and in real time

10.1. Potential errors
- Meaningless or irrelevant content
- Too long pauses, or pauses at a wrong place
- Inappropriate word or structure choice
- Strange intonation or pronunciation

The right form and structure
- Agreement errors
  - e.g. The report about the fires are very long
  - $<5\%$ in an error elicitation study (Bock & Miller 1991).
- Severe structural anomaly (unanalysable)
  - $0.5\%$ of utterances (Deese 1984).
- Sound/word errors (Garnham et al 1982):
  - Sound 3.2/10,000 word
  - Word 5.1/10,000 word

Speech errors: phonemes
Spoonerism: mixing up phonemes (Reverend Dr. William Spooner, 1844-1930).
A blushing crow. (crushing blow)
A well-boiled icicle. (well-oiled bicycle)
Is the bean dizzy? (dean busy)
The Lord is a shoving leopard. (a loving shepherd)
You have hissed all my mystery lectures. (missed...history)
You have tasted a whole worm. (wasted...term)
Three cheers for our queer old dean! (dear old queen, referring to Queen Victoria)

8 main speech error types:
These appear at all linguistic levels (phoneme, morpheme, word level).
Inducing speech errors experimentally?
Example:
Give Back/Get Book/Go Booth/Give Booth/Bad Goof → Respond

Anticipatory biasing
• dog bone
• dust ball
• dead bug
• doll bed
• barn door → „darn bore”

• White, white, white - What does the cow drink? milk
• poke, poke, poke - What is the white of egg called? yolk

Speech errors
1. reading list→leading list (C anticipation)
2. a phonological rule→ a phonological fool (C perseveration)
3. brake fluid→blake fluid (C cluster division and C exchange)
4. speech error→ peech error (C deletion)
5. box of flowers→blocks of flowers (C anticipation and insertion; C cluster division)
6. fill the pool→ fool the pill (V exchange)
7. Sue weeded the garden→see weeded the garden (V anticipation)
8. annotated bibliography→ annotated bablibography (V perseveration)
9. drop a bomb→ bop a dromb (consonant cluster-C exchange)
10. when you get older your spine shrinks→ your shrine spinks
   (consonant cluster exchange).
• → C, V and C cluster are units of speech production.
• anticipation, perseveration, exchange

Are most basic units of speech production sub phonemic distinctive features?
Many errors involve phonological features:
V-features are mixed up with other V-features, C-features with C-features, but not C and V features.
1. big and fat→pig and vat; Is Pat a girl?→ Is bat a curl? (voicing reversal)
2. cedars of Lebanon→cedars of Lemadon (nasality reversal)
3. he’s a vile person→ he’s a file person (voicing anticipation)

Speech errors are not random
1. Phoneme exchange in similar positions
   – *a hissed mystery lectures (missed history lectures)
   – *b burst of beaden (beast of burden)
   – *c nife lite (night life)
2. Phonemic environment is important too (see c)
3. Cs for Cs, Vs for Vs
4. Resulting forms follow the phonological rules of the language (slick + slippery _slickery,
   no phonologically illicit forms like slickperry or slipkery)
5. Consistent intonation patterns in speech errors
Errors in normal production
- Speech errors are common in normal production as well, characteristically
  - Segment substitution within words
    - ‘cuff of coppee’, ‘cussy pat’
  - Phoneme perseveration
    - put a those puppies need a blanket
  - Blending words or syllables
    - How long have you played the clariola?
    - Irregardless
  - At the sentence level
  - Errors are syntactically consistent (nouns for nouns, verbs for verbs)
  - Generally they preserve the intonation contour of the intended utterance
  - Suggests that syntax, prosody and meaning is assembled first, and words pop into their
    place rather late in the process
  - Typical sentence-level errors
  - Exchanges
    - Data was stored on a laboratory in our computer
  - Antonym substitution
    - It’s near the front...back door
  - Feature mixup
    - What’s that big light up there?
    - The image is too loud up close to the screen

Summary:
1. We only insert phonemes from our own language
2. Exchanges follow the rules of syntax, even if they are physically or semantically far apart
3. Intonation contours are built before uttering the words

Speech errors: what do they suggest?
Bock & Levelt (1989)
1. The message is planned according to meaning
2. Lexical selection (ensures that words are available, but does not access them)
3. Assignment of function
4. Positional processing – assembly of components (word order) & suffixing (‘s, ‘ing etc)
5. Prosodical & phonological planning
6. Access of words
7. Uttering the sentence
   - i.e. after selecting the word and at the very end of the articulation process
     (after prosody etc.)
  - Most of the time we only realize we made a word-level error after we produced it.

![Diagram of Garrett's model of speech production]

<table>
<thead>
<tr>
<th>Stage</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Identification of meaning—a meaning to be conveyed is generated.</td>
</tr>
<tr>
<td>2</td>
<td>Selection of a syntactic structure—a syntactic outline of the sentence is constructed, with word slots specified.</td>
</tr>
<tr>
<td>3</td>
<td>Generation of intonation contour—the stress values of different word slots are assigned.</td>
</tr>
<tr>
<td>4</td>
<td>Insertion of content words—appropriate nouns, verbs, and adjectives are retrieved from the lexicon and placed into word slots.</td>
</tr>
<tr>
<td>5</td>
<td>Formation of affixes and function words—function words (articles, conjunctions, prepositions), prefixes, and suffixes are added.</td>
</tr>
<tr>
<td>6</td>
<td>Specification of phonetic segments—the sentence is expressed in terms of phonetic segments, according to phonological rules.</td>
</tr>
</tbody>
</table>

An example of how we produce an utterance based on Garrett’s (1975, 1976) model of speech production:

(A) **Message level**
   Intention to convey particular meaning activates appropriate propositions

(B) **Functional level**
   SUBJECT = “mother concept”, VERB = “wipe concept”, OBJECT = “plate concept”
   TIME = past
   NUMBER OF OBJECTS = MANY

(C) **Syntactic frame**
   (DETERMINER) N1 V [+PAST] (DETERMINER) N2 [+PLURAL]

(D) **Lexical access**
   /mother/ /wipe/ /plate/

(E) **Positional level**
   (DETERMINER) DETERMINER) /mother/ /wipe/+[PAST] (DETERMINER) /plate/ + [PLURAL]

(F) **Function words are phonologically specified**
   /the/ /mother/ /wiped/ /the/ /plates/

(G) **Low-level phonological processing and articulation**

**Syntactic planning**

Functional level:
- No word order
- Generation of syntactic frame
- Selection of phonological representation of content words from lexicon

Positional level:
- Insertion of content words into syntactic frame
- Selection of phonological form of content words

**Evaluation of Garrett’s model**

- Words can exchange even across great distance, because we retrieve them before determining their positions
- Sounds are only defined after positional level, which constrains phoneme exchanges (short distance)
- Word exchanges in the same syntactic category, securing insertion of words into syntactic frame
- Late phonological specification of function words blends, cognitive intrusions are not explained, because it is a serial model, and these suggests parallel processing

Blend: it’s difficult to validate (validate+verify)

Get out of the clark - instead of car, seeing Clark’s above a shop (high level processes constrained by low-level ones as phonological similarity)
Syntactic planning
- It seems there is syntactic priming in real life
- Branigan, Pickering, & Cleland (2000): speakers tend to reuse their own/other speaker’s syntactic constructions
  - Speakers tend to use the constructions they just read (Potter & Lombardi, 1998)
  - Speakers tend to use their own/other speaker’s syntactic constructions (Branigan, Pickering, & Cleland, 2000)
- Classical design
- Prime (hear a sentence) → target (picture description)

**Bock (1986): study of syntactic ‘persistence’ with picture description**
.a: The ghost sold the werewolf a flower  
.b: The ghost sold a flower to the werewolf

**Bock (1989): global syntactic role counts, syntactic priming does not depend on lexical similarity**

**Bock, Loebell & Morey (1992)**
- Priming experiment:
  - Goal: The alarm clock woke up the boy/The boy was woken by the alarm clock
- Syntactic priming:
  - More active descriptions after active primes
    - Five people carried the boat
    - The alarm clock woke the boy
  - After passive primes more passive descriptions:
    - The boat was carried by five people

**Bock et al (2)**
- Conceptual priming:
  - Tendency to put an entity with the same animacy in subject position again:
  - Inanimate subject of a passive sentence
    - Subject of an active sentence is inanimate too
    - The boat was carried by five people → The alarm clock woke the boy
Syntactic and conceptual priming
  • Conceptual factors influence grammatical encoding
  • What is the explanation?
    – It affects assignment of grammatical function
      A functional processing effect
    – It affects word order
      A positional processing effect

10.3. From thought to speech

How does a mental concept get turned into a spoken utterance?

Levett, 1989, 4 stages of production:
1. Conceptualising: we conceptualise what we wish to communicate (“mentalese”).
2. Formulating: we formulate what we want to say into a linguistic plan.
   – Lexicalisation
   – Lemma Selection
   – Lexeme (or Phonological Form) Selection
   – Syntactic Planning
3. Articulating: we execute the plan through muscles in the vocal tract.
4. Self-monitoring: we monitor our speech to assess whether it is what we intended to say,
   and how we intended to say it.

Levett’s (1989) “Blueprint for the Speaker”

[Diagram of Levett’s 1989 model of speech production]

Psycholinguistik – Folien, WS11

Seite 130
From Thought to speech:

- Numerous evidence for a two-stage level of lexical access
  - Tip-of-the-tongue phenomenon
  - Semantic and phonological errors appear at different stages of production
- Arguments for/against interaction are less equivocal
  - Data on the time course of activation suggest NO interaction
  - Modeling speech errors suggests interaction

Two levels of lexical access
- Original model: from semantics directly to word form
- In most models, there is an intermediate stage called ‘lemma’ level
- A lemma is an abstract word form – it contains semantic and syntactic information
- In certain languages (like French or Italian), it also specifies gender

Temporal evidence for the 2-stage model
stage 1: conceptual => lemma

Levrlt et al. (1991): EARLY PRIMING: KANGEROO => FROG
**Picture-word interference**

- Word naming, disregarding the associated word
  - semantically: picture of a dog, *cat*
  - phonologically: picture of a dog, *dot*
  - No relation: picture of a dog, *cup*
- Schrieffers et al. (1990)
  - Changes in word-picture timing
- Word first $\Rightarrow$ semantic inhibition
- Picture first $\Rightarrow$ phonological facilitation
  - semantic processing first, phonological later
  - With early presentation of distractors, only semantic distractors had an effect
  (slower picture naming)
  - With late presentation of distractors, only phonological distractors had an effect
- Similar results for a lexical decision task too (Levelt, 1990)
- Suggests that processing has different stages
- But there are arguments for interaction as well
• The lemma connects meaning and form.
• Solves the hyperonym problem (dog/animal). The lemma gives semantic info as a whole, not as a set of features
• This semantic representation then specifies the phonological form of the word
• Two-stage access is manifest in the ‘tip-of-the-tongue’ phenomenon.
Overview

• Activation of the concept TIGER (x)
  – A certain amount of activation to the concept LION as well
• This activates the lemma tiger
  – And the nodes noun, feminine, etc
• Selection of the word tiger (in the normal case)
• Selection of the word form tiger
  – And all of its phonemes, /t/, /r/, etc.
  – articulation

Levčet et al.’s model

• modular
  – Only one lemma is selected (and sends activation to the next level)
  – No feedback during encoding
• E.g. phonology cannot affect lemma selection
• Clear distinction between syntax/semantics and phonology
• Other models question several elements of the model, but most assume a lemma level
  (except Caramazza)

Two stages provide a good explanation

1. **Semantic and phonological errors**
   
   • Semantic errors: \( \text{cat} \rightarrow \text{dog} \): we access the lemma, but not the phonological form
   
   • Phonological errors \( \text{cat} \rightarrow \text{hat} \): at selection of phonological form

   ☹ mixed errors \( \text{cat} \rightarrow \text{rat} \)

   ☹ blends \( \text{children} \ + \ \text{young} \rightarrow \text{chung} \)

2. **ToT**
   
   ☹ partial phonological information
“Suppose we try to recall a forgotten name. The state of our consciousness is peculiar. There is a gap therein: but no mere gap. It is a gap that is intensely active. A sort of wraith of the name is in it, beckoning us in a given direction, making us at moments tingle with the sense of our closeness, and then letting us sink back without the longed-for term. [...] The rhythm of a lost word may be there without a sound to clothe it; or the evanescent sense of something which is the initial vowel or consonant may mock us fitfully, without growing more distinct.”

William James

The tip-of-the tongue phenomenon

- What is the navigational instrument used to measure the angle distance between any two visible objects, especially the height of the sun, moon and stars at see?
- Often phonologically similar words come into mind as approximations of the target word (secant, sextett, sexton)
- In French or Italian, the gender of the word is often known as well (e.g. Vigliocco et al, 1997; Badecker et al, 1995)
- They are better than chance in guessing the countable/uncountable dimension of the noun (Vigliocco et al., 1999)
- In a two-stage model this is explained by the failure of accessing the phonological form with successful access of the lemma level
• Vigliocco et al. (1997)
  – Subjects presented with word definitions
  • Gender was always arbitrary
  – If unable to retrieve word, they answered
    • How well do you think you know the word?
    • Guess the gender
    • Guess the number of syllables
    • Guess as many letters and positions as possible
    • Report any word that comes to mind
  – Then presented with target word
    • Do you know this word?
    • Is this the word you were thinking of?

  – Scoring
    • + TOT
      – Both reported some correct information in questionnaire
      – And said yes to recognition question
    • - TOT
      – Otherwise
  – Results
    • + TOT: 84% correct gender guess
    • - TOT: 53% correct gender guess
      – chance level
  – Conclusion
    • Subjects often know grammatical gender information even when they have no phonological information
    • Supports split between syntax and phonology in production

Caramazza, 1997
Dispenses with lemma level, unnecessary complication
Sometimes at least partial phonological info is available, when gender information is not: it should not be possible to retrieve phonological info without syntactic
10.4. Spreading Activation Theory (Dell)

**Dell et al.’s (1997) Interactive activation model**

- four levels of activity
  - Semantic (meaning)
  - Syntactic (grammatical structure of words in the planned sentence)
  - Morphological (basic units of meaning or word forms)
  - Phonological (sounds)
- representation formed at each level
- processing occurs simultaneously at all levels
- uses speech errors as primary data
- Lexicon: connectionist network containing nodes for concepts, words, morphemes, and phonemes
- Insertion rules (which is highest activated?) determine items selected for insertion into sentences
- Errors predicted by model:
  - Errors more likely when speaker has not formed a coherent speech plan
  - Errors should be from same category
  - Anticipation errors (because of multiple activations; “The sky is in the sky”)
  - Exchange errors (because once selected, items’ activation turns to zero) (“I hit the bat with my ball”)
- Dell et al.’s model is an interactive activation model (1986, 1997, 2000) in which activation cascades from one level to the next
- If the target word is ‘CAT’, the semantic level activates the lemma of cat and all the related nodes (pl. DOG, RAT)
- All relevant phonological units – of the target word and of competing words (neighbors) too get some activation during lemma selection
- Resting activation level is determined by the frequency of activation – frequent words need less extra activation to be selected
• Feedback from phonological units means that irrelevant lemmas are activated too—MAT, LOG, HAT
• Feedback updates activation of previous levels
• Activation from different levels is summed up, until one word gets enough activation to be selected.
• Selection points (extra activation shots to the most active node) introduce seriality and weaken the effect of competing candidates
• Errors occur in noise, when a nontarget word gets higher activation than a target word
• Noise has several sources e.g. frequency (more errors on less frequent words), distraction, tiredness, brain damage
• The model produces errors naturally, but it can also be ‘damaged’ to simulate brain damage.
• Damage can be simulated in two ways: making connection weights smaller or making information degradation faster

10.5. Weaver++ (Levelt, Roelofs et al., 1999)

• Word Encoding by Activation and Verification
• It does not actually simulate errors, in a two stage model, errors can occur at different levels
• When we access the lemma of a word, errors are going to be semantically related words
  – lorry → bus
• Phonological errors sound similar to the target word:
  – soldier → shoulder
• These occur when we access the phonological form of the word
• These errors are evidence that there are two levels of lexical access

Phonological encoding

• When the word form is accessed it needs to be phonologically encoded
• Weaver++ assumes different levels for morphological encoding, syllabification, and phonological encoding
• Phonological encoding involves combination of phonological and syllable information
• A sound might be inserted in the wrong place yielding a nonword
• Substitutions: anymay (anyway)
• Exchanges: par cark (car park)
• Insertions: plublicity (publicity)
• Deletions: dugs (drugs)

Stages of WEAVER++

- **Conceptual preparation**: potential lexical concepts are activated. E.g. if you wished to communicate to someone that some trees fell, the concepts of oak, tree, wind and fall might be activated.
- **Lexical selection**: an abstract word (lemma) is selected, along with its syntactic features. E.g. the lemmas corresponding to tree and fall might be selected, together with the fact that tree is a noun and fall is a verb.
- **Morphological encoding**: the basic word form of the selected lemma is activated. For example, if the lemma is tree, then morphological encoding will activate the plural word “trees”. If the lemma is fall, this will be activated in the past tense form, “fell”.
- **Phonological encoding**: the syllables of the word are computed.
- **Phonetic encoding**: speech sounds are prepared using a syllabary (i.e. a dictionary of commonly used syllables in the language).
- **Articulation**: the actual physical utterance of the word

Discrete versus interactive processing

- Models differ as to how activation flows from one level to the other
- Weaver++ is a discrete or sequential model – processing at one level has to be completely finished before getting to the next level
- The model is feed-forward, i.e. processing on one level influences processing on the level below it, but not vice versa.

- Items related to each other at one level are activated in parallel (e.g. dog/cat/rat)
- A single node is selected when it gets enough activation
- Competitive words do not send activation to the next level
- No interaction (overlap) between levels
Predictions of and Evidence for the WEAVER++ model

- The model posits that all the syllables of a word to be uttered are prepared before the word is articulated.
- The model therefore predicts that it should take longer to begin saying a word with many syllables than a word with few syllables.
- Levelt (1999): participants were asked to name items which had either one or two syllable names in a timed experiment.
- The results were in agreement with the WEAVER++ model: people took longer to begin saying the two syllable names than the one syllable name.
- The WEAVER++ model explains the tip of the tongue phenomenon as correct functioning of the conceptual preparation, lexical selection and morphological encoding stages, followed by failure at the phonological encoding stage.
- Studies of brain-damaged patients also provide empirical support for the WEAVER++ model.
11. PRÜFUNGSFRAGEN

11.1. Beispiele Fragen

- Wie lässt sich die „broca aphasie“ beschreiben?
- Was ist „syntactic priming“?
- Nenne 4 Merkmale von “child-directed speech”!
- Was ist der Wada-Test?
- Was ist „two-sentence“?
- Beschreibe die „size of neighbourhood“!
- Was ist die „contrast theory“?
- Was ist „lemma“?
- Erkläre „garden-path“ und was man damit forscht!
- Erkläre Logogen’s „Bootstrapping model“ der Semantik!
- Erkläre „categorical perception“!
- Erkläre das „gavagi problem“!
- Erkläre „double association“!
- Erkläre “surface vs. deep structure“!
- Give 4 factors influencing sentence processing!
- Erkläre Berko’s “Wug-Test“!
- Describe minimal attachment and give an illustration!
- Erkläre “taxonomic constraint“!
- Erkläre “syntactic bootstrapping“!
- Erkläre „uniqueness point“!
- What is cohort?
- Erkläre “poverty of stimulus“!
- Erkläre “dichotronic listening task“!
- What is a linguistic universal? Give an example.
- What is a lemma? Give an example.
- Describe semantic/syntactic bootstrapping.
- Give an example of a garden-path-sentence.
- What is a critical period?
- What is the principal of a late closure? Give an illustration.

11.2. Beispiele Essays

- Autonomy of syntax
- Innate language acquisition
- Human and animal speech
- Speech production
- Nativist model
- 1 model of language production in detail
- Autonomy of the syntax and ambiguous sentences
- Human vs. animal communication-give 5 examples of differences!