

# HS466: INTRODUCTION TO LINGUISTICS

VAIJAYANTHI M. SARMA

E-mail: vsarma@iitb.ac.in

Office hours: H&SS, by appointment

Moodle: For all announcements, posting of reading material, discussions, grading etc.

## Course information

Attendance: As per the rules of IITB you are required to be in class!

Evaluation:

- Unannounced, open-notes, in-class assignments, 5 marks each (20 marks)
- Mid-term exam (30 marks)
- One extended problem set (10 marks)
- 1 final exam (40 marks)

Gayathri G. is the TA for the course.

## Course information

Textbook:

- (1) Linguistics: An Introduction to Language and Communication, 5th Ed, MIT Press by Akmajian, Adrian, Demers, Richard A., Farmer, Ann K., Harnish, Robert M. Prentice-Hall India, Low-priced edition
- (2) Introduction to language (7th ed) by Fromkin, Victoria, Robert Rodman and Nina Hyams. Thomson, 2003

Other readings: As provided in class/on moodle

When we study human language, we are approaching what some might call the “human essence”, the distinctive qualities of mind that are, so far as we know, unique to man.

-Noam Chomsky

## What is Linguistics?

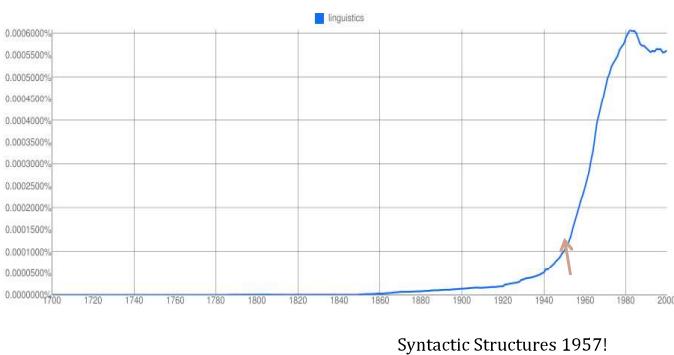
Linguistics is the scientific study of language, in particular, some aspects of human language and its structure.

Some facets of language make it apparently unique in the biological world and in the study of cognition. Moreover, language is *creative* and *complex* in a way that warrants careful study.

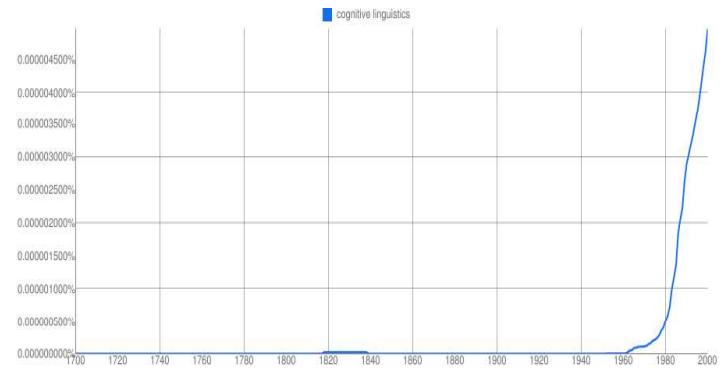
## Linguistics is concerned with questions such as

- How do languages work? Are there rules? What are rules?
- What do we know when we know a language?
- What's the range of possible human languages?
- How do people use language? Are there conventions about that?
- What are the biological and developmental properties?

## Google N-Gram for “linguistics”



Syntactic Structures 1957!



## What is Language?

- Linguistics is a (relatively recently constructed) scientific discipline with established theories, analytic methods, and real-world applications.
- The primary object of linguistic study is *human language*, not language in other extended senses or claims about how language ‘should’ be, made by so-called experts or language mavens.
- Linguistics is the study of **language**, not individual languages. By language we mean the system of principles that account for linguistic expressions; languages that actually exist (or existed) and are used by people in different speech communities.

- For most linguists, **language** (from Fr. *langue* “tongue”) is the (underlying, implicit) **pattern** of human speech, and the systems that enable us to speak and to listen to our language(s).
- Other phenomena are also called “language” because of close connections, similarities or analogies to this central aspect:

writing  
sign languages  
computer languages and codes  
the language of dolphins or bees

The ordinary-language meaning of the word reflects this process of semantic extension from a speech-related core. The etymology of the word, makes the same point.

## Definition From the American Heritage Dictionary

[lan-gwāj] (NOUN).

1. a. The use by human beings of voice sounds, and often of written symbols that represent these sounds, in organized combinations and patterns to express and communicate thoughts and feelings.  
b. A system of words formed from such combinations and patterns, used by the people of a particular country or by a group of people with a shared history or set of traditions.
2. A nonverbal method of communicating ideas, as by a system of signs, symbols, or gestures: “the language of algebra.”
3. Body language.
4. The special vocabulary and usages of a scientific, professional, or other group.
5. A characteristic style of speech or writing: “ribald language.”
6. a. Abusive, violent, or profane utterance: “language that would make your hair curl (W.S. Gilbert).”  
b. A particular manner of utterance: “gentle language.”
7. The manner or means of communication between living creatures other than humans: “the language of dolphins.”
8. Language as a subject of study.
9. The wording of a legal document or statute as distinct from the spirit.
10. Computer Science. Machine language.

- The phenomena named by the extended senses are quite different from one another.
- Writing is a system of transcription for speech.
- Deaf sign languages are an expression in a different medium of the same underlying human capabilities as spoken language.
- Computer languages are artificial systems with some formal analogies (of debatable significance) to the systems underlying human speech.

## Discipline

Some linguists think that the boundary between the patterns of spoken language and other modes of communication is not a sharp one, or even that it is entirely artificial. For them, the extended senses of the word "language" belong to the same subject matter as the core sense.

A larger proportion of poets, philosophers and religious thinkers agree with them, often going on to view language as magically connected to the world it describes: *In the beginning was the Word, and the Word was with God, and the Word was God.* ... (John 1:1, Bible)

The core of the field of linguistics has always been the analysis of linguistic structure, and this course will introduce the basic concepts of this disciplinary core. There is much intellectual, practical and human interest in other aspects of the study of language, and we'll survey these too.

### How to annoy a linguist? ... Just say:

- Oh, you're a linguist! How many languages do you speak?

- Some linguists are polyglots; most aren't.
- Polyglot: Etymology, Greek *polyglōttos*, from *poly-* many + *glōtta* language; speaking or writing several languages : multilingual

### The Linguist's Motto

Linguistics is descriptive, not prescriptive.

### Prescriptive Grammar

- Rules against certain usages.
- Few if any rules for what is allowed.
- Condemns forms generally in use.
- Explicitly normative enterprise.

## Kinds of Prescriptive Rules, "Correctness"

- Established criteria of educated written language, e.g. third-person singular /s/: she goes, not she go.
- Issues on which educated people differ (and which may be different in written and spoken forms), e.g. who/whom did you see? Or, the data is/are unreliable
- Changes in the spoken language that some people resist: e.g. between you and I/me; you and I/me can do this together
- Pure inventions of self-appointed grammarians with no basis in actual usage, e.g. prohibition of dangling prepositions, split infinitives or It is I

## Some prescriptive rules

- Don't split infinitives (e.g. "...to boldly go where no man has gone before")
- Don't use double negation (e.g. "I didn't do nothing")
- Don't end a sentence with a preposition (e.g. "What is she talking about?")
- Don't use who in place of whom (e.g. "Who did you see?")
- Don't misuse hopefully (e.g. "Hopefully he will arrive tomorrow")

From Dear Abby, April 9, 2002  
<http://www.uexpress.com/dearabby>

It has been nearly 10 years since these rules of basic grammar appeared in my column – and we can all use this refresher course. My pet peeve – double negatives: I don't know nothing and We don't go nowhere are the worst offenders.

## Prescriptivist propaganda

Using two negatives in a sentence to convey a negative is illogical (two negatives make a positive). So you're not supposed to say 'You don't owe me nothing' to mean 'you are free of debts to me'. But many languages employ just this syntax, using multiple negative words to convey negation. These include even the most literate forms of some languages:

Personne n'a rien dit (French)  
Nobody not-has nothing said  
'Nobody said anything.'

## Anyway, language isn't logical

- parkway vs. driveway
- maternity dress vs. paternity suit
- bathing trunks (pl) vs. bikini (sing)
- you are vs. \*you is
- Aren't I clever? vs. \*I aren't clever
- \*Amn't I clever? vs. I am clever

## Sample Prescriptive Rule

Don't end a sentence with a preposition.

Dear Dr. Science,  
Is there a reason for not ending a sentence with a preposition that you can think of?  
–John Mostrom from Seattle, WA

## Dr. Science's Reply

I must admit I don't know where you're coming from. Correct usage in English and Science is something I've devoted my whole life to. Of course, if I say anything you can't understand, it will just be a new hammer you can try to hit me or another expert over the head with. There are plenty of people like you I can't hope to change the mind of. But then, I've dealt with people like you before. People who don't really want to learn, but just hope to find someone they can publicly disagree with. There is little I can say that your type won't find something to object to. But getting back to your question; no, there's really no reason for not ending a sentence with a preposition, at least none I can think of.

- After an overzealous editor attempted to rearrange one of Winston Churchill's sentences to avoid ending it in a preposition, the Prime Minister scribbled a single sentence in reply:
- "This is the sort of arrant nonsense up with which I will not put."

## Standard Complaint

- 'Decline' of English: For centuries, they have been claiming that 'incorrect' English will surely lead society into disarray and even violence.
- "If we allow standards to slip to a stage where good English is no better than bad English, where people turn up filthy ... at school ... all those things tend to cause people to have no standards at all, and once you lose standards there's no imperative to stay out of crime."  
--British Government Minister, 1985
- "Tongues, like governments, have a natural tendency to degeneration."  
--Samuel Johnson, 1775

## Alan Ross

U	Non-U
looking-glass	mirror
scent	perfume
have a bath	take a bath
sick	ill
rich	wealthy
wireless	radio
What?	Pardon?

## Language as gatekeeper

- ... "they said, "All right, say 'Shibboleth'." If he said, "Sibboleth" because he could not pronounce the word correctly, they seized him and killed him at the fords of the Jordan (River). Forty-two thousand Ephraimites were killed at that time by Gileadites. (Hebrew Bible, Judges 12)
- A modern parallel took place in the Dominican Republic in 1937, when thousands of Haitians were massacred on the basis of whether or not they could roll the /r/ in the Spanish word for "parsley", *pejeril*. (President Rafael Trujillo)

## Where do prescriptive rules come from?

When we look to find the origin of some well-known prescriptive rules of English, we find that they have some rather surprising sources.

- Rules adopted into English from Latin
- Rules adopted from mathematical logic
- Speech patterns imposed by speakers with high social prestige
- Word choices of older generation

## Status of prescriptive rules

- \*"Quien ha hablado con?"  
(literally: *who did he speak with* is terrible Spanish)
- Infinitives in Romance languages = 1 word  
comere: *to eat* (Latin)  
andare: *to go* (Italian)  
estar: *to be* (Spanish)  
danser: *to dance* (French)
- English  
"I didn't see nothing."  
"He didn't never say nothing like that."
- *negative concord or negative agreement:*  
Spanish: "No vi nada." = *I didn't see nothing.*  
French: "Il n'a jamais dit cela." = *He hasn't never said that.*
- Ya, right.

- Decision to follow is an individual choice
- Can substantially affect a listener's judgment of a speaker
- Associated with prestige and education
- No scientific basis for its prestige

## Descriptive Grammar

- Rules characterizing what people do say.
- Goal to characterize all and only what speakers actually do.
- Caveat: Competence vs. Performance.
- Tries to do so in a way that reflects internalized generalizations,
- What people know about their language.

## Where did you learn grammar?

[in a room full of boys]

Would you use he/his, they/their or something else?

Everyone insisted that \_\_\_\_ answer was correct.

## Where did you learn grammar?

[in a room full of boys]

Everyone drives \_\_\_\_ own car to work.

## Where did you learn grammar?

[in a room full of boys]

Everyone was happy because \_\_\_\_ passed the test.

## Where did you learn grammar?

[in a room full of boys]

Everyone left the room, didn't \_\_\_\_?

## Where did you learn grammar?

[in a room full of boys]

Everyone left early. \_\_\_\_ seemed happy to go home.

## Descriptive Linguistics

Who taught you this? Where did you learn it?

- Not in school. Not (all) from your parents.
- From exposure, you learned a system. That system is grammar in the linguist's sense.

- All of language is rule-governed.
- Linguists try to describe those rules precisely.
- A set of those rules is what linguists call a *grammar*.
- Linguists theorize about the properties of *natural grammars*.

## Some Descriptive Rules of English

- The subject precedes the verb, the object follows the verb  
"Wallace fed Gromit"  
*not* \*"Fed Gromit Wallace" (order in Hindi)
- Auxiliary verbs precede the subject in questions  
"What has she done?"  
*not* \*"What she has done?
- Form the plural of a noun by adding '-s'  
"three books"  
*not* \*\*"three book"

## Correlates

- **all languages have grammar**, Warlbiri vs. English  
The two dogs now see several kangaroos  
Dogs two—now—see—kangaroos several
- **all grammars are equal**, ability to express any proposition – equal as instruments of communication or thought
- **all grammars change** – over time, new vocabulary, new syntax, new meaning
- **universality of grammar** - set of grammars learned and used by humans is small, 'learnable' grammars
- **tacitness of grammatical knowledge**

## Where do descriptive rules come from?

- Saying what 'makes sense'? No.
- Saying what is easily understandable? No.
- 'Facts of life' with no communicative function – seems so.

### One rule that does not aid comprehension

- The subject precedes the object (in English)  
The dog bit the boy.  
The boy bit the dog.
- Absent in German - case markers show subject and object:  
Der Hund hat den Jungen gebissen.  
Den Jungen hat der Hund gebissen.

## Likewise

- The object follows the verb.  
I know that the dog ate the cookie.  
I know that the dog the cookie ate.

- The verb agrees with the subject.  
Wallace likes Wendoline.  
Wallace like Wendoline.

## Some constructions disallow certain verbs

The millionaire gave the museum a painting.  
The millionaire sent the museum a painting.  
The millionaire bought the museum a painting.  
The millionaire offered the museum a painting.  
The millionaire obtained the museum a painting.  
The millionaire donated the museum a painting.  
The millionaire presented the museum a painting.

- Comparative '-er'  
big -> bigger  
sad -> sadder  
red -> redder  
happy -> happier  
but not  
enormous -> enormous  
melancholy -> melancholy  
crimson -> crimson  
delighted -> delighted

## Hard to Understand (but perfectly grammatical!)

The students who knew all the answers were sleeping in their rooms when the rest of the class was taking the test.  
The doctor who the intern who the nurse supervised had bothered lost the medical reports.

That we lost bothers me.  
That that we lost bothers me seems to upset her.  
**That that that we lost bothers me seems to upset her isn't surprising.**

## Status of Descriptive Grammar

- Not a personal choice whether or not to follow it, it's automatic
- Doesn't require explicit teaching (for children)
- Social prestige is irrelevant
- "Ungrammatical" is different from "Incomprehensible"

## Some Linguistic Methods

- Fieldwork
- Formal analysis of patterns in data sets
- Psycholinguistic experiments
- Computational modeling
- Corpus analysis

## Summarizing

We are going to investigate the idea that language is innate or an 'instinct' by

(a) Looking at the structure of language; and

(b) Looking at how language works in the brain, it is acquired, develops in history, compares with communication in other species, etc.

## Some myths and facts about language

Language means controversies!

### Some questions are entirely political:

- Should governments try to accommodate speakers of minority languages?
- How important is it to maintain rigorous standards of usage?
- Is it bad to borrow words from other languages rather than inventing native ones? Should we weed out non-native words?

Other questions are factual, though they have immediate practical consequences:

- Does bilingual education work?
- What are the consequences of oral education for deaf children?
- To what extent can ordinary citizens understand legal contracts?
- How well do computer speech recognition systems work?

## But the myths...

### A third set of questions interests those who care about language itself:

- Are Korean and Japanese derived from the same historical source?
- How much of linguistic structure is innate, and how much emerges from the experience of communication?
- Why will most English speakers delete "that" in "this is the book [that] Kim told me about," but not in "this is the book [that] impressed Kim so much"?

**Myth:** speech and writing are parallel forms of linguistic expression, different but equally fundamental types of text.

**Fact:** Speech is primary, writing is secondary and is always derivative of speech.

**Myth:** Non-standard dialects are degraded and error-ridden versions of standard languages.  
**Fact:** Standard languages are either an arbitrary choice among a range of geographical and social dialects, or an artificial construct combining aspects of several dialect sources. Ways of speaking that happen not to be "standardized" in this way have their own history, at least equally valid even if lacking in prestige.

**Myth:** Primitive cultures have primitive languages, at a lower level of development and less well able to express a wide range of ideas.

**Fact:** There are no primitive languages; there are no demonstrated differences in fundamental communicative efficacy among languages. (Though obviously the range of available vocabulary varies greatly.)

**Myth:** Women speak about twice as fast as men, and use about three times as many words per day.

**Fact:** There is a great deal of individual and situational variation in how much people talk and how fast they talk. Many studies show no meaningful difference in averages between males and females. No studies appear to support the story about the greater gabbiness (or superior communicative prowess) of women.

A husband looking through the paper came upon a study that said women use more words than men. Excited to prove to his wife that he had been right all along when he accused her of talking too much, he showed her the study results. It read "Men use about 15,000 words per day, but women use 30,000".

The wife thought for a while, then finally she said to her husband "It's because we have to repeat everything we say."

The husband said "What?"

## Introduction contd.

### Some Areas of Linguistics

- Phonetics: The study of speech sounds
- Phonology: The study of sound systems
- Morphology: The study of word structure
- Syntax: The study of sentence structure
- Semantics: The study of linguistic meaning
- Pragmatics: The study of language in use

These six categories divide up the study of the linguistic system itself. Each category focuses on a different *level of description and analysis*.

Speech communication depends on conventional connections between sound and meaning. To understand how it works, we need to describe and analyze the sounds, the meanings, and the words and phrases that connect them.

### Phonetics

### Phonology

The study of speech sounds

- Voicing: [s] versus [z]
- sue ([su:w]) versus zoo ([zu:w])

The study of sound systems

house (sing. noun) ([haws])  
 house (verb) ([hawz])  
 houses (plural noun) ([hawsɪz])

### American vs. British English

bar	ba(r)
leader	leade(r)
carton	ca(r) ton
corner	co(r) ne(r)
carry	carry
hairy	hairy
Britain	Britain
ride	ride

### Standard vs. Lhasa Tibetan

gàbàr	gàbà(r)	'where?'
phósár	phósá(r)	'young man'
márpó	má(r)pó	'red'
sérná	sé(r) ná	'greedy'
gàrè	gàrè	'what?'
chörà	chörà	'debating hall in a monastery'
rìnpóché	rìnpóché	'precious jewel, monk'
ràngwáng	ràngwáng	'freedom'

## Morphology

The study of word structure

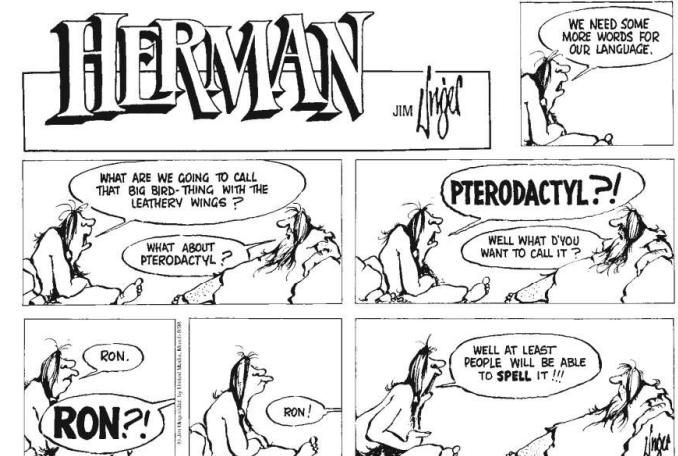
- missile: 'ICBM'
- anti-tank-missile: 'missile targetting tanks'
- anti-aircraft-missile: 'missile targetting aircraft'
- anti-missile-missile: 'missile targetting ICBMs'

## Morphological Rules

- Rule: Anti-X-missile is a missile targetting Xs.  
What kind of missile targets anti-missile-missiles?
- anti-[anti-missile-missile]-missile
- anti-[anti-[anti-missile-missile]-missile]-missile:  
'missile targetting anti-anti-missile-missile-missiles'

## More word building

- industry
- industri-al
- industri-al-ize      'having to do with industry'  
'cause to have to do with industry'
- industri-al-iz-ation      'the process of causing things to have to do with industry'
- industri-al-iz-ation-al      'related to the process of causing...'
- industri-al-iz-ation-al-ize      'cause to be related to the process of causing things...'
- But \*industry-ize \*industri-al-ation
- cats kids horses wug-? bez-? xerox?



"Herman"® is reprinted with permission from Laughing Stock Licensing Inc., Ottawa, Canada. All rights reserved.

## Syntax

The study of sentence structure

- The policeman saw the thief with the telescope.
- I forgot how good beer tastes.
- Visiting relatives can be dangerous

- The woman will sit on the chair.
- The general will urge on the troops.
- Which chair will the woman sit on?
- Which troops will the general urge on?
- On which chair will the woman sit?
- On which troops will the general urge?
- The woman will sit the chair on.
- The general will urge the troops on.

## Competence vs. Performance

- Why did she say that George should resign?
- Why did she **not** say that George should resign?
  
- John thinks that George is a genius
- Mary said that John thinks that George is a genius
- Susan can't believe that Mary said that John thinks that George is a genius

Some sentences conform to the rules of grammar, but put extreme demands on our processing abilities.

- That we lost bothers me.
- That that we lost bothers me seems to upset her.
- That that that we lost bothers me seems to upset her  
isn't surprising.

Chomsky: Distinguish the abstract system of rules (the competence grammar) from the reality of language use (performance).

## Semantics

"What's one and one?" "I don't know," said Alice. "I lost count." "She can't do Addition," the Red Queen interrupted.

LEWIS CARROLL, *Through the Looking-Glass*, 1871

The study of linguistic meaning

- We screened the candidates.
- I saw her duck.
- [huwz an fəst]  
Who's on first?/Hu's on first.

## Pragmatics

- I didn't tell her that Mary was very smart.  
(*her=Mary?*)
- I didn't tell her mother that Mary was very smart.  
(*her =Mary?*)
- She didn't say that Mary left.  
(Mary may have left, or not)
- She didn't regret that Mary left  
(Mary has to have left)

The study of language use

- Q: Is Bush a Republican?  
A: Is the Pope Catholic?
  
- Can you close the window?  
(really means "Close the window!")
- Did you know that there's a bear behind you?  
(really means "Look out!")

## Other areas of study

- **Historical Linguistics:** How languages change over time.
- **Sociolinguistics:** How languages vary socially. How language is used as a social resource.
- **Areal linguistics:** Languages of the world, typology
- **Psycholinguistics:** What goes on in people's heads as they use language.
- **Neurolinguistics:** The biology of language, the hardwiring in the brain
- **Language Acquisition:** How people learn language (first language acquisition; second language acquisition).
- **Computational Linguistics:** Making computers process (generate/'understand'/translate...) human languages.

These categories tell us about possible connections between linguistics and external topics. Each of this second set of linguistic sub-disciplines can in principle deal with any of the six levels of description in the first set.

- Thus sociolinguists study the social dimensions of pronunciation (phonetics or phonology), word and sentence structure (morphology and syntax), conversational styles (pragmatics), and so on.
- Psycholinguists have studied perception, production and learning of a similar range of topics.

## Levels of grammar and Historical Linguistics

- **phonology:** IE \*kʷekʷlos 'wheel' > Gk. *kuklos*, Skrt. *chakra*, Eng. *wheel*
- **syntax:** Speaks he English? > Does he speak English?
- Thou (you) madest (made) me
- **semantics:** Old English *fæger* 'fit, suitable'; Modern English *fair* came to mean 'pleasant, enjoyable' then 'beautiful' and 'pleasant in conduct', from which the second modern meaning 'just, impartial' derives. The first meaning continued to develop in the sense of 'of light complexion' and a third one arose from 'pleasant' in a somewhat pejorative sense, meaning 'average, mediocre', e.g. *He only got a fair result in his exam.*

## More lexical meaning

- **Lexical change**
  - *Gentle* was borrowed in Middle English in the sense of 'born of a good-family, with a higher social standing' (gentleman). Later the sense 'courteous' and then 'kind, mild in manners' developed because these qualities were regarded as qualities of the upper classes.
  - *Lewd* (Old English *læwede*) originally meant 'non-ecclesiastical, lay', then 'uneducated, unlearned', then to 'vulgar, lower-class' and then via 'bad-mannered, ignorant', to 'sexually insinuating'.
  - *Sophisticated* meant 'unnatural, contaminated' but now has the sense of 'urbane, discriminating'. Cf. *sophistry* (from Old French *sophistrie*) still has its original meaning of 'specious, fallacious reasoning'.

## Next

- The list of topics related to language or language use is open-ended, and so the second list could be extended almost indefinitely (forensic linguistics, dialectology, language education, metrics, and so on).

- Phonetics

## Languages of the world

- Diversity of languages
- Language grouping & language universals
- Language death and revival

### Diversity of Languages

### UNESCO data

- World population: 7 billion people (2012)
- Number of languages spoken (estimated): 6000+ languages
- Over a third of the world's population speaks just one of 8 languages!
- 389 languages are spoken by 94% of the population, and the rest are spoken by the remaining 6%!
- 2200 languages in Asia and 260 in Europe and 800 in Papua New Guinea
- About 2200 languages are in UNESCO's endangered list
- About 40 languages have only 1 speaker left!
- The greatest diversity loss is in India, Brazil and USA

	0 - 9999	10000 - 99999	100000 - more	No data on number of speakers	Total
Vulnerable	337	149	104	5	595
Definitely endangered	407	151	62	24	644
Severely endangered	431	50	10	38	529
Critically endangered	509	15	0	50	574
Extinct	220	0	0	9	229
Total	1904	365	176	126	2571

### Language Grouping

Some widely spoken Asian languages

Mandarin 937 mi  
(14%)

Hindi 295 mi

Bengali 189 mi

Arabic 442 mi

1863 mi

Some widely spoken European languages

English 322 mi

Spanish 332 mi

Russian 170 mi

Portuguese 170 mi

994 mi

- Ways of **grouping** languages

1. Common history
2. Common sounds or grammar (typology)

## Historical grouping

English	German	French
foot	Fuss	pied
hand	Hand	main
blood	Blut	sang
drink	trinken	boire
full	voll	plein
stone	Stein	perre

## Government & administration terms

English	German	French
government	Regierung	gouvernement
prince	Furst	prince
court	Koningshof	court
society	Gesellschaft	société
Pray	beten	prier
judge	Richter	juge

## Accidental similarity

Language A	Language B	English
ban	Bhanem	'woman'
allaban	Alnoba	'person'
lion-obhair	lhab	'netting'
dun	odana	'town'
claden	kladen	'snowflake'
bata	pados	'boat'
cuithe	cuiche	'gorge'
monadh	monaden	'mountain'

- Language A is Scots Gaelic (Scotland). Language B is Northeastern Algonquian (Maine).
- Language classification relies on
  - (a) the **Uniformitarian Principle**, that knowledge of processes that operated in the past can be inferred by observing ongoing processes in the present (i.e. languages work now they way they did before) and
  - (b) (b) **regularity of sound change**, that sound change is regular and exceptionless.

## Regular sound changes

- Grimm's Law (von Schelgel, Rasmus Rask and Jakob Grimm): Systematic comparison between languages; first Germanic sound shift of inherited PIE stops into Germanic

Sanskrit	Greek	Latin	Gothic	English	PIE
pita	pate:r	pate:r	fadar	father	*p̥eter-
padam	poda	pedem	fotu	foot	*ped-
bhratar	p̥hrate:r	frate:r	broþar	brother	*bhrater-
bharami	p̥hero	fero	baira	bear	*bher-
sanah	hene:	senex	sinista	[senile]	*sen-
				three	*trei-
deka	dekem	taihun	ten		*dekm-
he-katon	kentum	hund(rab)	hundred		*dkm-tom-

Change	Germanic (shifted) examples	Non-Germanic (unshifted) cognates
*p → f	English: foot, Dutch: voet, German: Fuß, Gothic: fótu, Icelandic, Faroese: fótur, Danish: fod, Norwegian, Swedish: fot	Ancient Greek: πόδις (pódis), Latin: pēs, pedis, Sanskrit: पदा, Russian: под (pod), Lithuanian: pėda, Latvian: pēda
*t → þ [θ]	English: third, Old High German: thriht, Gothic: þridja, Icelandic: þridji	Ancient Greek: τρίτος (trítos), Latin: tertius, Gaelic: treas, Irish: tri, Sanskrit: त्रेता, Russian: третий (tretyj), Lithuanian: trečias, Latvian: trīs
*k → h [x]	English: hound, Dutch: hond, German: Hund, Gothic: hunds, Icelandic, Faroese: hundur, Danish, Norwegian, Swedish: hund	Ancient Greek: κύων (kyón), Latin: canis, Gaelic: cù, Welsh: ci
*kʷ → hw	English: what, Gothic: hwa ("who"), Danish: hved, Icelandic: hveð, Faroese: hvat, Norwegian: hva	Latin: quod, Gaelic: ciod, Irish: cad, Sanskrit: क्वा-, kim, Russian: кво- (ko-), Lithuanian: ką, Latvian: kas
*b → p	English: warp, Swedish: värpa, Dutch: weper, Icelandic, Faroese: varpa, Gothic: wairpan	Latin: verber
*d → t	English: ten, Dutch: tién, Gothic: taihun, Icelandic: tju, Faroese: tiggju, Danish, Norwegian: ti; Swedish: tio	Latin: decem, Greek: δέκα (déka), Gaelic, Irish: deich, Sanskrit: दस्, Russian: десять (desyat'), Lithuanian: dešint, Welsh: deg, Latvian: desmit
*g → k	English: cold, Dutch: koud, German: kat, Icelandic: káldur, Danish: kold, Norwegian: kold, Swedish: kall	Latin: gelū, Russian: холод (kholod)
*gʷʰ → kw	English: quick, Frisian: quick, queck, Dutch: kwiek, Gothic: qius, Old Norse: kvíkr, Danish: kvik, Icelandic, Faroese: kvíkr, Swedish: kvick, Norwegian: kvikk	Lithuanian: gyves
*bʰ → b	English: brother, Dutch: broeder, German: Bruder, Gothic: broþer, Icelandic: bróðir, Danish: brødr, Swedish, Norwegian: broder	Sanskrit: भ्रातर् (bhātr), Russian: брат (brat), Lithuanian: bros, Old Church Slavonic: братъ (brat''), Latvian: brālis
*dʰ → d	English: door, Frisian: deur, Dutch: deur, Gothic: dæw̥r, Icelandic: dyr, Faroese: dyr, Danish, Norwegian: dør, Swedish: dörr	Irish: doras, Sanskrit: द्वार (dvera), Lithuanian: durys, Latvian: durvis
*g → g	English: goose, Frisian: guos, Dutch: gans, German: Gäns, Icelandic: gæs, Faroese: gás, Danish, Norwegian, Swedish: gås	Russian: гусь (gus')
*gʷʰ → gw	English: wife, Proto-Germanic: wibaz (from former gwibaz), Old Saxon, Old Frisian: wib, Old High German: wib, German: Weib, Old Norse: víf, Icelandic: víf, Faroese: vív, Danish, Swedish, Norwegian: vív	Tocharian A: kip, B: kwipe (vulva)

## Grammatical Grouping: by Phonology

- a. Clicks: Bushman languages of Kalahari Desert (S. Africa), Damin, N. Australia
- b. Front rounded vowels: Includes Romance languages, Germanic languages, and Finnish, and excludes some German-speaking areas, i.e. not a historically defined set of languages.
- c. Syllable typology: CV, CVC, CCVCC etc.

## Grammatical Grouping: by Morphology

- **Isolating** languages: one word, one form

Khi toi den nha ban toi, chung toi bat dau lam bai.  
when I come house friend I, plural I begin do lesson.  
'When I came to my friend's house, we began to do  
lessons.' (Vietnamese)

- **Synthetic/Inflecting** languages: prefixes, suffixes, stem changes express meanings
  - trink-en 'drink' (German)  
drink-1prs.plur.pres
  - sing-s (-3prs.sing.pres)
  - sang (-past)

- **Polysynthetic/Incorporating** languages: words built of many pieces: stems and inflections
  - anngyaghllangyugtuq
  - anngya- ghlla- ng- Yug- tuq
  - boat- AUGMENT-acquire-want-3.sing
  - 'He wants to acquire a big boat'
  - (Siberian Yupik, Eskimo)

## Grammatical Grouping: by Syntax

### word order

- SVO The dog chased the cat (English, French, Thai)
- SOV The dog the cat chased (Japanese, Eskimo, Turkish)
- VSO Chased the dog the cat (Irish, Hebrew, Tagalog)
- VOS Chased the cat the dog (Huave, Coeur d'Alene; Hadza, Tanzania, Malagasy, Madagascar)

## Syntactic Generalizations/universals

- "If a language has SOV word order, then auxiliary verbs follow the verb"
- English: SVO, verb follows auxiliary
- German, Japanese, Korean: SOV basic word order, verb precedes auxiliary

## Language Death and Survival

- "The coming century will see the death or the doom of 90% of mankind's languages", Michael Krauss, Alaskan Languages Center
- Fact: Two-thirds of the world's languages have less than 10,000 speakers
- **Some causes of language death (different from 'dead' language)**  
few speakers, no young speakers (no intergenerational transmission), low prestige for language, contact with other languages (assimilation), linguicide

## Examples

- **Martuthunira** (NW Australia): speakers killed off by exploitation and disease ; **Yiddish** (Central Europe): speakers killed or dispersed by Holocaust ; **Faroese** (North Atlantic); **Native American** languages: influence of English
- **Andamanese languages**: Aka-Bea, Akar-Bale, Aka-Kede, Aka-Kol, Oko-Juwoi, A-Pucikwar, Aka-Cari, Aka-Kora, Aka-Jeru; 36 speakers in 1997, bilingual in Hindi, Aka-Bo, Onge; 96 speakers in 1997, mostly monolingual, Jarawa; 200 speakers in 1997, monolingual.
- The **Ahom language**, also called Tai-Ahom, spoken by the Ahom people who ruled most of Assam from 1228 to 1826. The language today is used chiefly for liturgical purposes, and is no longer used in daily life. (Wikipedia).

## Language revival & creation

- young speakers, increased prestige, lack of contact
- Hebrew, creation of new territory in earlier part of this century (the only real example). Work on Nicaraguan languages like Ulwa (Misumalpan) etc.

## Language change

- Diachronic change, the example of Negative agreement in French

### Old French: No Negative Agreement

- a. Il ne mange  
*He NOT eats*  
'He doesn't eat'
- b. Il ne marche  
*He NOT walks*  
'He doesn't walk'
- c. Il ne boit  
*He NOT drinks*  
'He doesn't drink'

### Reinforcing negation in Old French

- a. Il ne mange **mie**  
*He NOT eats crumb*  
'He doesn't eat a crumb'
- b. Il ne marche **pas**  
*He NOT walks step*  
'He doesn't walk a step'
- c. Il ne boit **goutte**  
*He NOT drinks drop*  
'He doesn't drink a drop'

### ■ Reinforcing negation in English (‘Negative Polarity Items’)

- a. He doesn't eat.
- b. He doesn't eat at all
- b'. \*He eats at all
- c. He doesn't eat a crumb
- c'. ≠He eats a crumb

### ■ Contemporary ‘formal’ French

- a. Il ne mange **pas**  
*He NE eats NOT*  
‘He doesn't eat’
- b. Il ne marche **pas**  
*He NE walks NOT*  
‘He doesn't walk’
- c. Il ne boit **pas**  
*He NE drinks NOT*  
‘He doesn't drink’

### ■ Double Negatives in Old English

Ic ne can noht singan  
*I not can nothing sing*  
‘I can't sing anything’  
[noht = nothing, nought]

- The reinforcement with ‘pas’ has been reanalyzed as negation, with the result that there is negative agreement in standard French. In English ‘ne’ disappeared and ‘noht’ came to be reanalyzed as the entire negation. The spoken variety of French has now moved towards dropping the *ne* and keeping the original emphaser alone!

### ■ 1200 (similar to French)

ich ne seye not  
he ne speketh nawt

- 1400 *ne* → Ø,  
I seye not the words  
we saw not the knights

- Aux inversion with negation (no main verb)  
I did not say the words  
(slow development in Middle English)

## Standards: Pidgins and creoles

which develop rapidly among speakers who need a new common language -- for instance:

- Haitian Creole (6+ million speakers in Haiti and New York City)
- Tok Pisin (2 million speakers in Papua New Guinea)
- Jamaican Creole or Patois (2 million speakers)
- Hawaiian Creole (1/2 million speakers)
- Palenquero (3,000 speakers in Colombia)

## Unwritten languages

or languages where writing is hardly ever used -- whose form is set by spoken interaction only:

- Ilocano (5.3 million speakers, Philippines)
- Chagga (800,000 speakers, Tanzania)
- Buang (10,000 speakers, Papua New Guinea)
- Toda (1000 speakers)

## Written languages with no academies

-- for instance  
English (322 million speakers)  
Marathi (70 million speakers)

## Languages with academies

- French (109 million speakers; academy established 1635)
- Spanish (266 million speakers; academy established 1713)
- Hungarian (14.4 million speakers; academy established 1830)
- Hebrew (2.7 million speakers; academy established 1953)

Languages codified to preserve an archaic form, for instance:

- Latin
- Old Church Slavonic
- Sanskrit
- Tamil

## The Language Instinct

- We see that speakers 'know' rules, which permit them to create and understand an infinite number of sentences. But how do children learn these rules? If there is no explicit teaching going on (and there isn't) how do they 'infer' the 'correct' rules from the necessarily limited data they are exposed to? Possible hypotheses:
  - (a) an all-purpose learning mechanism, or
  - (b) a language specific system.
- Hypothesis (b) suggests that there exists a **Universal Grammar**, which is **innate**, is the basis for **all human languages** and permits the **acquisition of language**.
- [Set of human languages ↔ Generated by UG]

## Universal Grammar

- We are all born knowing (tacitly) certain things about how human language is put together. This is how we manage to acquire this very complex system, without relevant instruction; we have a head start.
- Linguistics is the study of the aspects of the human mind that allow us to acquire and use language.

## Examples of instincts:

- Bats use sonar to echolocate; homing pigeons know where home is; deer rub antlers against trees; dolphins play; cats wash themselves in a characteristic manner; cuckoos' song etc.
- Special properties of individual species, not related to 'general intelligence' of the species or the individual
- Instincts develop (more or less) automatically, often needing relatively little input from parents or other members of the species

## Infinite Use of Finite Means

- Speakers 'know' a finite number of rules
  - How can we 'know' an infinite number of sentences?
- A simple rule:** If S is a sentence of English, *Nobody cares that S is a sentence of English.*
- John is asleep
  - Nobody cares that John is asleep
  - Nobody cares that nobody cares that John is asleep
  - Nobody cares that nobody cares that nobody cares that John is asleep
  - etc.
- Simple iteration of the rule to sentence (a) yields an infinite number of Ss

## The Problem of Induction

- A rule, applied either explicitly or implicitly, is required to make generalizations
- Example 1: Given a series of numbers {3, 9, 12, 15, 21}, determine a common property of these numbers and which of these two numbers 17 or 18 belong to this group.

■ **Example 2:** Question formation in English.

- a. John is in the garden
- b. Is John \_\_ in the garden?
- c. John is in the garden next to someone who is asleep
- d. Is John \_\_ in the garden next to someone who is asleep? (First *is*)
- e. \*Is John is in the garden right next to someone who \_\_ asleep? (Second *is*)

Possible rule of Q-formation: move to the beginning of the S the first *is* which is uttered.

- f. A unicorn that is eating a flower is in the garden
- g. \*Is a unicorn that \_\_ eating a flower is in the garden? (First *is*)
- h. Is a unicorn that is eating a flower \_\_ in the garden? (Second *is*)

■ The child must have something to guide his acquisition of a rule that goes beyond the sentences that he has heard.

## Arguments in favor of the Language Instinct

### Theory B

## Two possible hypotheses

- about the acquisition of human language
- **Theory A:** Acquisition of language is guided by an all-purpose learning mechanism (for example, general intelligence)
  - **Theory B:** Acquisition of language is guided by a language specific mechanism that drives language learning alone. This is Universal Grammar. (Innate design underlying the grammars of all human languages, guides the child with respect to what is possible, what is not. It is a modular system.)

## (1) Uniformity Throughout Human Species

- Similarities between the world's languages
- If the UG hypothesis is true, then all the languages of the world should have common structure(s), since they must all obey the rules of that instinct/UG. This is the main concern of modern linguistics which aims to uncover the underlying principles/parameters/design of UG. We will see evidence when we deal with actual structures.
  - All humans master a human language, except in certain extreme circumstances. This makes language a characteristic property of the species (unlike other things that are unique to humans, but that only a few humans can master, e.g. chess, tap-dancing, piano)

## 'Poverty of the Stimulus'

- No explicit teaching, learning through observation and induction.
- Data are impoverished, simple and not sufficient to determine the complex rules acquired by children.

## (2) Species specificity



- Other species simply can't learn human language. Despite some fairly heroic attempts to teach other species human language, no one has managed to bring chimps to a level remotely approaching the sophistication of human language. Question: Should they? Should we attempt to swing from trees? To echolocate?
- The communication systems of other animals are not structured in the way human language is, do not use the same (UG) design properties. These systems typically encode a very small, fixed set of expressions, such as bees signaling direction, quantity etc. of pollen, vervet monkeys sounding alarm calls etc. Human language is characterized by remarkable *rule-governed creativity*.

### (3) The invention of language by children

- What is the outcome in the absence of coherent linguistic input? Do children acquire a language? Is it different or undetermined? Do they 'create' a language and is that language in conformity with UG and does it reveal facts about UG?

### Compare the following

- No, the men, ah-pau [finished] work – they go, make garden. Plant this, ah cabbage, like that. Plant potato, like that. And then – all that one – all right, sit down. Make lilly bit story.
- When work pau [is finished] da guys they stay go make [are going to make] garden for plant potato an' cabbage an' after little while they go sit down an' talk story.

### (3) contd:

#### Example 1: Creoles and Pidgins

- **Creoles** are (spoken or sign) languages created by children who grew up in a pidgin-speaking environment. (LI 20 sugarcane plantations in Hawaii with Phillipino, Korean, Portugese, Chinese, Japanese and Puerto Rican labourers).
- **Pidgins** develop when people speaking different languages must communicate; presence of a dominant/colonial language (French, Dutch, English, etc.)
  - mostly content words
  - few function words, no case, agreement, tense
  - use of simple sentences – no embedding
  - little/inconsistent syntax, or syntax 'borrowed' from native languages (Hawaiian pidgin, V-final (Japanese), V-initial (Filipino))
  - all pidgin users have some other 'native' language

- Now days, ah, house, ah, inside, washi clothes machine get, no? Before time, ah, no more, see? And then pipe no more, water pipe no more.
- Those days bin get [there were] no more washing machine, no more pipe water like get [there is] inside house nowadays, ah?

- **Creoles** show the properties of a full-fledged language:
  - uses word order, embedded clauses (Hawaiian Creole:SVO)
  - vocabulary from dominant language
  - all creoles similar, even though based on different languages
  - details of grammar, tense, determiners (indefinite and definite articles), relative clauses, questions, word order
- Yiddish (German + Hebrew), Modern Hebrew (from an artificially simplified form of ancient Hebrew), Modern German (many German-speaking tribes), Japanese (a Korean-like language + an Austronesian language?), ASL, Hindi?

- Building - high place - wall part - time - nowtime - and then - now temperature every time give you.
- Get one [there is an] electric sign high up on da wall of da building show you what time an' temperature get [it is] right now.

- [the > da, a > wan, bin = past]
- 'John went to Honolulu to see Mary'  
(outcome is not grammatically encoded in English – he may/may not have seen Mary)
- John bin go Honolulu go see Mary (he did see Mary)
- John bin go Honolulu for see Mary (he failed to see Mary)]

### Example 3: Spontaneous 'errors'

- Don't giggle me! (MAKE ME GIGGLE)
- Go me to the bathroom before you go to bed (MAKE ME GO)
- The tiger will come and eat David and then he will be died (KILL)
- I want you take me a camel ride over your shoulders into my room (GIVE)
- Be a hand up your nose (PUT)
- You can push her mouth open to drink her (FEED)  
(c.f. melt, bounce, race, break)

- First schools for the deaf formed by the Sandinista government (1979), taught lip-reading and speaking. Sign language was discouraged, developed in the schoolyard.
- **Lenguaje de Signos Nicaraguense (LSN)** – use with friends, families by 17-15 year olds (developed when 10 years and older). Suggestive elaborate circumlocutions. No inflections.
- **Idioma de Signos Nicaraguense (ISN)** – developed by young children exposed to LSN: Compact, stylized, inflections, fluid signing

Special parental input is not necessary. All children learn their native language(s) if they are simply put into an environment where they receive plenty of exposure to the language. In fact, young children are not particularly good at following such instructions anyway:

- Child: Nobody don't like me.
- Adult: No, say 'nobody likes me'.
- Child: Nobody don't like me... *8 repetitions*
- Adult: No, now listen carefully. Say, 'Nobody likes me'.
- Child: Oh! Nobody don't likes me.

### Example 2: Invention of Language, Deaf Isolates

- 'Simon' was studied by Prof. Elissa Newport of the University of Rochester. When deaf children receive imperfect language input from hearing parents who have learned ASL as adults, the children have been observed to 'improve' upon the parental input, making it more consistent and more sophisticated. Simon sorted out the noise and used signs and inflections perfectly.
- Attempt at communication, eye contact
  - Gestures not performed on objects but in the air
  - Gestures in strings, separated by relaxing of the hand
  - Consistent use of word order, inflections etc.
  - Go beyond the here and now, point to objects that aren't present
  - Acquisition development 'parallels' speaking children – verbal and gestural babbling, 1 sign, 2 sign combinations, 3 signs etc.

Deaf children who receive no sign-language input from their parents have been found to spontaneously create simple 'languages' of their own, called 'Home Sign' (rudimentary language?).

### Caveat

- If there is a UG, then it is not surprising that children can 'invent' language. But this also follows from Theory A – the invention may be driven by the generalized learning strategies. Both theories posit some innate mechanisms – the question is, are these specifically linguistic mechanisms?

## (4) Double dissociations

between language and general intelligence

- Impaired language in people with normal intelligence and intact language in people with impaired intelligence establishes the independence between language ability and general abilities, suggesting that the rules of language are not entirely determined by general intelligence, but by a separate module (UG).

## Example 2a Impaired intelligence with normal language

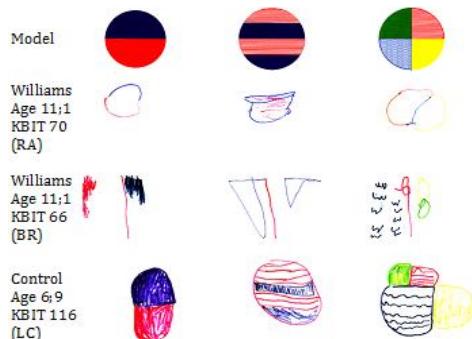
### Williams Syndrome

- Distinctive 'elf-like' facial appearance, a particular heart defect, abnormalities of other organ systems. Mild to moderate mental retardation. Severe deficits in spatial understanding, (navigation, copying see drawings), but verbal descriptions are rather good. Low general IQ. Poor math. Often good with music. Cannot tie shoe-laces, tell right from left, retrieve things from the cupboard, add two numbers, find their way, draw a bicycle. Understand complex sentences and fix ungrammaticalities. Fondness for unusual words. (Chromosome 7, ~26 genes are deleted)

## Example 1a normal intelligence with impaired language

### Broca's aphasia

- Broca's aphasics have normal general intelligence, but they have considerable production difficulties. Their comprehension, though apparently normal, is also imperfect – unable to use 'grammatical' information.
- Me...build-ing...chairs, no, no cab-in-ets. One,saw...then, cutting wood...working..."
- Cookie jar... fall over... chair... water... empty... ov... ov... [Examiner: 'overflow']
- The leopard killed the lion
- The leopard was killed by the lion



## Example 1b normal intelligence with impaired language

### Selective Language Impairment.

SLI patients have average IQs, but have linguistic difficulties with some aspects of morphology and syntax:

- It's a flying finches, they are.
- She remembered when she hurts herself the other day.
- The neighbors phone the ambulance because the man fall off the tree.
- The boys eat four cookie.
- Carol is cry in the church.

## Crystal says...

- And what an elephant is, it is one of the animals. And what the elephant does, it lives in the jungle. It can also live in the zoo. And what it has, it has long, gray ears, fan years, ears that can blow in the wind. It has a long trunk that can pick up grass or pick up hay...If they are in a bad mood, it can be terrible...If the elephant gets mad, it could stomp; it could charge. Sometimes elephants can charge like a bull can charge. (Ursula Bellugi)

## **Example 2b Impaired intelligence with normal language**

- **Hydrocephalics** (caused by spina bifida, idiot savants, chatterbox syndrome). Split spine, increase in cerebrospinal fluid within the ventricles of the brain, cortex thin. Severely retarded, cannot read or write.
- They've got a habit of doing that. They lose, they've lost my bank book twice, in a month, and I think I'll scream. My mum went yesterday to the bank for me. She said, "They've lost your bank book again." I went, "Can I scream?" and I went, she went "Yes go on." So I hollered. But it is annoying when they do things like that. TSB trustees aren't...uh the best ones to be with actually. They're hopeless. (Denyse, R. Cromer)

## **(5) Critical period for language learning**

- Victor, 'the wild boy of Aveyron' 1799: few words, 11-12 years. Taught by Jean-Marc Gaspard Itard. Categorization. No syntax.
- Isabelle, 1940s, 6 years. Cognitive skills of 2 year old. Normal.
- Chelsea, profoundly deaf, 1980s, 31 years, 2000 words. Normal cognition.
- 2<sup>nd</sup> language learners and 'accents'

## **Conclusions**

- Language is specific to humans, and uniform among humans
- Humans create language without instruction
- Language abilities are partly independent of other cognitive abilities
- Language learning requires a young brain
- Language has the properties of an 'instinct'.

# Phonetics

The Sounds of Speech

[fənetɪks]

[fʊnɛtɪks]

[fnerɪks]

## Phonetics

Inventory and structure of phones – the sounds of speech.

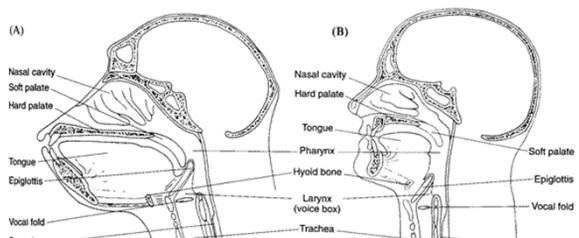
- ❑ Not all sounds that can be made by the human vocal tract are linguistic (hum, cough, moan, scream, whistle).
- ❑ There is not an infinite number of sounds that can be made by the human vocal tract. Each language (arbitrarily) selects/uses a subset of the possible phones.
- ❑ The phones are 'unambiguously' represented in the International Phonetic Alphabet which uses symbols and diacritics to capture the phonetic value of each sound.

## Specialization of speech

- Speech is an overlaid function on anatomy used for other primary purposes
- | Anatomy     | Primary function                 | Speech function         |
|-------------|----------------------------------|-------------------------|
| Lungs       | CO <sub>2</sub> , O <sub>2</sub> | air for speech (source) |
| Vocal folds | seal over bronchus               | voicing                 |
| Tongue      | move food                        | Vs and Cs               |
| Teeth       | chew, bite                       | articulation            |
| Lips        | seal oral cavity                 | articulation            |
- Voice box descends around 3-4 months, before that we have a primate vocal tract
  - More neural pathways – greater control over lips, tongue, palate
  - Vocal folds less fatty, more muscular
  - Speech breathing is slower; higher pressure, slower exhalation than in respiration

## The Speech Stream

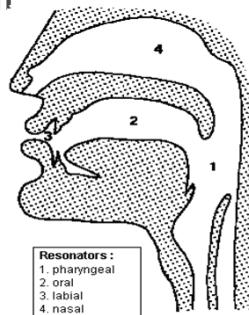
- Language comes out in a continuous speech stream, which we hear (if we speak the language) as a succession of discrete segments (phones/sounds). 'ball' : [bɔ:l]
- We can study the sounds of language from the perspective of their production (**articulatory phonetics**), of the physical result (**acoustic phonetics**), and of their reception and interpretation (**auditory phonetics**).



- Basic articulatory phonetics
  - How does the vocal tract work?
  - What kinds of sounds do languages use and how are they made?
- General principles of acoustic phonetics
  - What are the acoustic cues of language sounds?
  - What cues are utilized to classify speech sounds?

## Articulatory Phonetics: The Sound Producing System

- Resonators - cavities in the vocal tract that are used for speech



**Resonators:**

1. pharyngeal
2. oral
3. labial
4. nasal

## Anatomy of the larynx

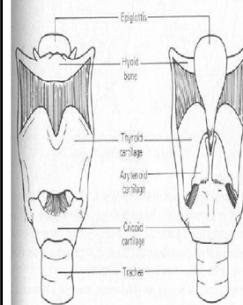
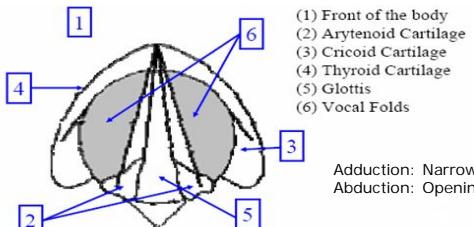


FIGURE 6.5.1 The larynx: anterior and posterior views

- **Larynx:** air from the lungs passes through the bronchi and the trachea through the 'voice box' or larynx or Adam's apple. The larynx is the opening of the trachea where it meets the pharynx. The larynx serves to close off the trachea during swallowing so that food is not admitted into the airway and facilitates swallowing by moving up against the back of the tongue.
- The main portion of the larynx is formed by the thyroid cartilage which forms the bulk of the larynx and which itself rests on the cricoid cartilage. The latter serves to reinforce the head of the trachea to keep the airway open. The flap-like epiglottis helps to shut off the airway during swallowing, as it swings down to meet the upraised larynx to prevent food from entering the trachea.

- Sound is produced when air is set in motion. The speech production mechanism consists of
- air supply (lungs)
  - a sound source that sets the air in motion (larynx)
  - and a set of filters or resonators (previous slide) that can modify the (source) sound in various ways



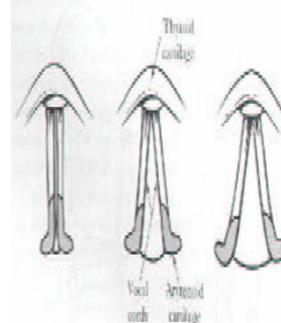
Adduction: Narrowing  
Abduction: Opening

- **Lungs:** air is expelled (egressive air stream) during speech for a majority of speech sounds. A certain level of pressure is maintained for a steady flow of air to keep the speech mechanism going steadily. The muscles involved are
  - the intercostals (between the ribs that compress the lungs slowly)
  - and the diaphragm (the sheet of muscle that separates the chest cavity from the abdomen).

Fine sheets of muscle flare from the inner sides of the larynx forming the paired vocal folds. Vocalization is enabled by manipulation of these cords/folds.

At the **posterior** end, the folds are attached to two triangular **arytenoid** cartilages which are themselves attached to the cricoid cartilage by means of the **cricoarytenoid** muscles. Flexing these muscles cause the arytenoids cartilages to pivot, causing the vocal folds to come together or separate. The space between the vocal folds is called the **Glottis**.

The vocal folds can be held tense or slack and open or closed. The larynx can also be raised and lowered.



13

(a) voicelessness (b) whisper  
 (c) modal voice (d) murmur  
 (e) falsetto (f) creak

- Voiceless – vocal folds pulled apart
- Voiced - vocal folds are close together but not tightly so, tense, periodic movements of folds
- Whisper – anterior portions are closed but the posterior regions are far apart
- Murmur –whispery voicedness
- Creaky – close vocal folds in low tension, very low frequency vibration
- Closed
- Falsetto, extension, high frequency, tense

## Articulatory phonetics

- **Speech sounds:** three kinds, consonants (complete or partial obstruction), vowels (little obstruction and sonorous) and semi-consonants/vowels or glides.
- **Articulators:** Passive and Active, Places of articulation
  - Tongue - tip, blade, body, back and root
  - Labial – bilabial and labiodental
  - Dental – interdental
  - Alveolar
  - Alveopalatal, palatal
  - Velar
  - Uvular
  - Pharyngeal
  - Glottal

14

## Glottal States

open: vf vf R front L  
 closed: vf vf R front L

Photo atlas of the University of Pittsburgh Voice Center at UPMC. Copyright 1998 UPMC Health System

## Parameters for Describing Consonants

- Place of articulation (previous slide)
- Glottal state (discussed earlier)
- Manner of articulation

### Manner of articulation:

- Oral vs nasal
- Stops/plosives/occlusives
- Fricatives, stridents and sibilants
- Affricates
- Liquids – laterals and rhotics
- Secondary: Aspiration, Labialization, Velarization etc.

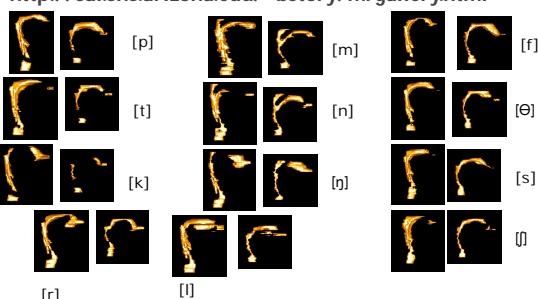
15

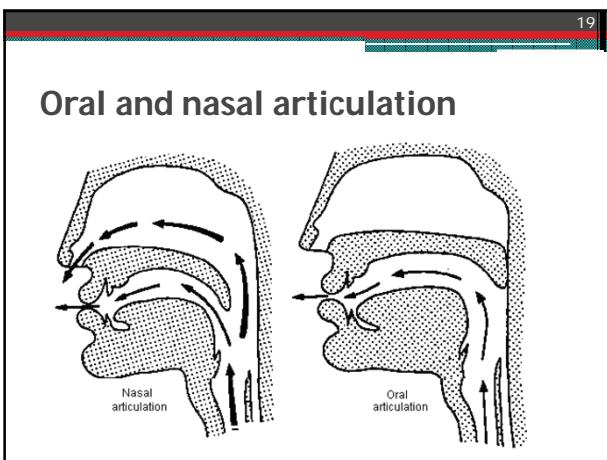
## The vocal tract, anatomy

Nasal cavity, Hard palate, Tongue, Pharynx, Hyoid bone, Larynx (voice box), Trachea, Oesophagus, Soft palate, Epiglottis, Vocal fold

## MRI of Vocal Tract

<http://sal.shs.arizona.edu/~bstory/mrgallery.html>





22

**Diacritics**

Syllabicity & Releases		Phonation	Primary Articulation	Secondary Articulation		
n	i	voiceless or slack voice	t b	voiceless or slack voice	t w d w	labialized
e	u	modal voice or strid voice	t d	apical	t d	palatalized
t b	t	(pre)aspirated	n a	breathy voice	t g d y	velarized
d		nasal release	n a	creaky voice	u t	Advanced
d l		lateral release			i f	Retracted
t		No audible release	n d	linguolabial	t z	Velarized or pharyngealized
e	ø	Lowered	ø	ø is a bilabial approximant	a j	Centralized
			e i	e i	ü	Mid-centralized
					e o	Retracted tongue root
						e o

20

### Consonants (pulmonic)

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Epi-glottal	Glottal
Nasal	m	nj		n		ɳ	ɲ	ŋ	ɳ			?
Plosive	p b		t d		t d	c ɟ	k ɡ	q ɣ			ʔ	ʔ
Fricative	f β	v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ɟ	x ɣ	χ ʁ	h	h	h
Approximant	w		r		l	j	ɥ					
Trill	B		r					R				
Tap or flap	v		t		l							
Lateral fricative		ɬ										
Lateral approximant		l		ɻ	ʎ							
Lateral flap		ɺ										

21

### Consonants (non-pulmonic)

	Anterior click releases (e.g. equi-posterior stops)	Voiced implosives	Ejectives
Bilabial fricated	b Bilabial	' Bilabial	examples:
Laminal alveolar fricated ("dental")	d Dental or alveolar	p' Bilabial	
Apical (post-alveolar abrupt ("retroflex"))	f Palatal	t' Dental or alveolar	
Laminal postalveolar abrupt ("palatal")	g Velar	k' Velar	
Lateral alveolar fricated ("lateral")	ç Uvular	s' Alveolar fricative	

**Consonants (co-articulated)**

M Voiceless labialized velar approximant

W Voiced labialized velar approximant

ɥ Voiced labialized palatal approximant

ʈ Voiceless palatalized postalveolar (alveolo-palatal) fricative

ɳ Voiced palatalized postalveolar (alveolo-palatal) fricative

ʃ simultaneous X and ʃ

Affricatives and double articulations can be joined by a tie-bar.

**C O T H E R E C O N S O N A N T S**

23

### Air Stream Mechanisms

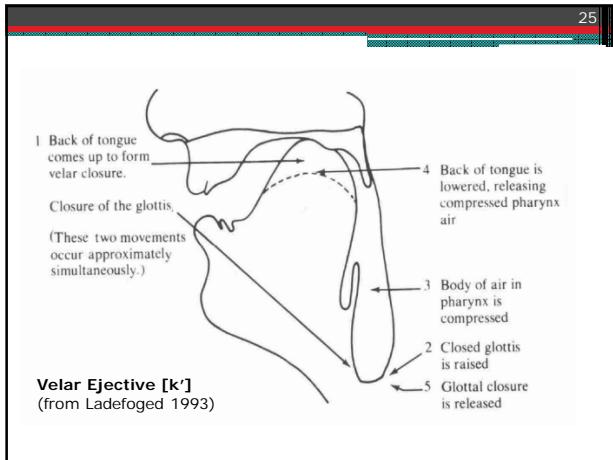
- Pulmonic Egressive**
- Pulmonic ingressive:** Speak while breathing in. Sounds odd, and is a little hard to control. This is not used regularly in languages, but you can do a *pulmonic ingressive uvular trill* if you try!

24

### Glottalic Air Stream Mechanism

- Glottalic egressive sounds**

**EJECTIVES:** /p', t', k'/. The glottis and velar port are closed, so that no air flows into or out of the lungs or nasal cavity. Then a stop-like occlusion is made. Pressure behind the stop closure is raised by raising the glottis, so that when the stop is released, there is a little plosive burst. Then the glottal stop is released. These sounds are voiceless because of the closed glottis.

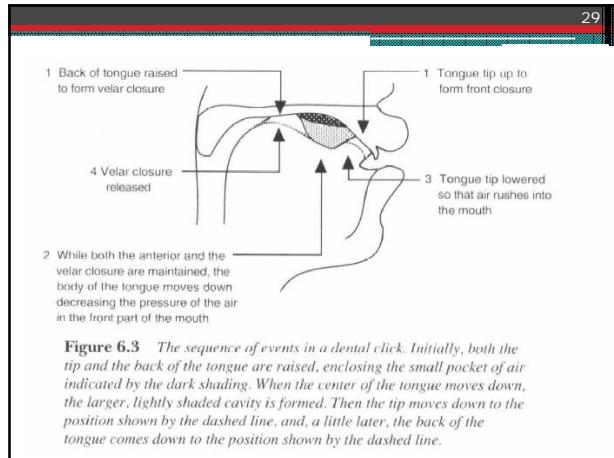
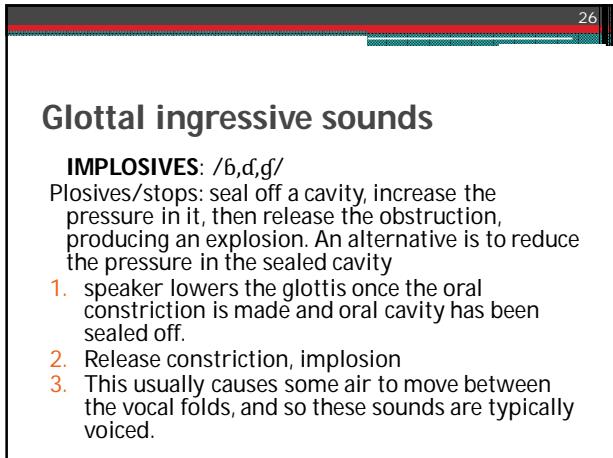


## Velaric Air Stream Mechanism

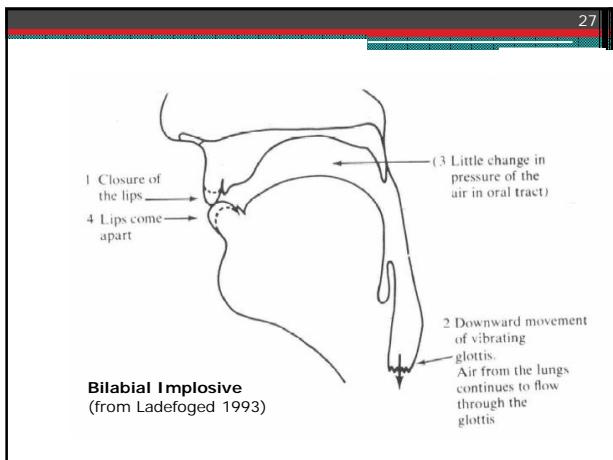
**Velaric ingressive sounds:** Clicks (sub-saharan languages)

1. A small cavity is sealed off in the mouth by raising the back of the tongue to the velum.
2. Another closure at the front with the front/blade of the tongue or with the lips. (bilabial, dental, alveolar, lateral etc.)
3. Then the jaw is lowered or the tongue depressed so as to create lowered pressure
4. Oral closure is released
5. Implosion on release
6. Release closure at velum

Only ingressive – egressive sounds not possible  
Possible to overlay nasalisation by using the glottis to nasal cavity route!



**Figure 6.3** The sequence of events in a dental click. Initially, both the tip and the back of the tongue are raised, enclosing the small pocket of air indicated by the dark shading. When the center of the tongue moves down, the larger, lightly shaded cavity is formed. Then the tip moves down to the position shown by the dashed line, and, a little later, the back of the tongue comes down to the position shown by the dashed line.



## DISTRIBUTION OF EJECTIVES

- 20% of the world's languages  
Quechua, Hausa, Amharic and Tigrinya, Nama and Sandawe, Itelmen etc.  
(NW America, Mexico, Guatemala, Caucasus, Ethiopia and neighbouring regions)  
The fictional language of the Na'vi in the film *Avatar*, distinguishes ejective consonants from plosives.

31

## DISTRIBUTION OF IMPLOSIVES

- about 13% of the world's languages

Found in many African languages, Vietnamese, in a few languages of the Amazon Basin, Mayan languages in North America, and Sindhi

34

## Diphthongs

The diagram shows a profile of a human head with arrows indicating the movement of the tongue tip and body during the articulation of various diphthong vowels. Labels include: *i*, *e*, *ε*, *ɔ*, *ə*, *ʌ*, *ɒ*, *ʊ*, *ɒ*, *ɒ*. The tongue tip moves from a low position at the beginning of the diphthong to a higher position at the end.

- The vowel quality changes during the course of the articulation of the diphthong. For example, in [ai], the tongue raises from a low to a high position and from a more central to a more front position in the oral cavity

32

## DISTRIBUTION OF CLICKS

- Clicks: Found only in Africa
- East Africa - Sandawe, Hadza (Tanzania, isolates); Dahalo, endangered Cushitic Language of Kenya
- South Africa: Khoisan
- Jul'hoan has 48 click consonants, among nearly as many non-click consonants, !Xóö and †Höä, Khoekhoe ("Nämä") of Namibia, !Kung
- Bantu
- Xhosa, Zulu, Sotho, Yeyi
- Damin, Lardil people, initiation language (only nasal clicks)

35

## MRI Vowels, M and F Vocal Tracts

<http://sal.shs.arizona.edu/~bstory/mrgallery.html>

A 4x4 grid of MRI scans showing the vocal tract configuration for various vowels. Each row contains four scans, each labeled with a vowel symbol and its name: *i, beat*; *ɛ, bit*; *e, bate*; *ε, bet* in the first row; *u, boot*; *ɔ, book*; *o, boat*; *ɔ, ball* in the second row; and *æ, back*; *a, barn* in the third and fourth rows respectively.

33

## Parameters for Describing Vowels

Vowels are produced without obstruction in the vocal tract. Vowel quality is determined by the resonating cavities created above all by the configuration of the tongue and lips. Vowels are described using the tongue location along the front-back and up-down axes. Lip (rounded or unrounded) and nasality are additional parameters. Length (duration) may also be a feature.

The diagram shows a profile of a human head with arrows indicating the movement of the tongue tip and body during the articulation of various vowel sounds. Labels include: *i*, *y*, *ɪ*, *ʊ*, *ɔ*, *ɒ*, *ə*, *ʌ*, *ɒ*, *ɒ*. The tongue tip moves from a low position at the beginning of the vowel to a higher position at the end.

36

## IPA Vowels

A chart of IPA vowels arranged in a diamond shape. The vertical axis represents height (Close at top, Open at bottom). The horizontal axis represents front-back position (Front on left, Back on right). The vowels are arranged as follows:

- Close: *i*, *y*, *ɪ*, *ʊ*, *ɔ*
- Close-mid: *e*, *ø*, *ɛ*, *θ*, *ø*
- Open-mid: *ɛ*, *æ*, *ʌ*, *ɒ*, *ɑ*
- Open: *a*, *æ*

Where symbols appear in pairs, the one to the right represents a rounded vowel.

Segmentals			Tone	37
Primary stress	Extra stress	Level tones	Contour-tone examples:	
Secondary stress [founə'tʃən]	é or ↗	Top	é or ↗	Rising
Long	e' Half-long	é ↗	High	ê ↘
Short	é Extra-short	ē ↗	Mid	é ↗
Syllable break — Linking (no break)	—	è ↗	Low	é ↗
ation		ë ↗	Bottom	é ↘
Minor (foot) break		Tone terracing	é ↘	Low falling
Major (intonation) break	↑	Upstep	é ↗	Peaking
Global rise	↓	Downstep	é ↘	Dipping
Global fall				

Hindi Consonants		40
क ka [kə]	ख kha [kʰə]	ग ga [gə]
च ca [tə]	छ cha [tʰə]	ज ja [dʒə]
ट ta [tə]	ঢ় ঢ়া [tə]	ঢ় ঢ়া [dʒə]
ত ta [tə]	ঢ় ঢ়া [tə]	ঢ় ঢ়া [dʒə]
প pa [pə]	ঢ় প়া [pə]	ব ba [bə]
য ya [jə]	ৰ রা [rə]	ল la [lə]
শ sha [ʃə]	ঢ় শা [ʃə]	ব্ল ব্লা [və]
হ ha [hə]		
Additional consonants (only used in loanwords)		
ক্ষ	খ্স খ্সা গ্স জ্স জ্সা	ঢ্ক্স ঢ্ক্সা ফ্স ফ্সা
Common conjunct consonants		
ক্ষা ksa	জ্জা জ্জা	ত্ত্বা ত্ত্বা
ধ্ব dda	ন্ন ন্না	ঢ্ড ঢ্ডা ঘ্ব ঘ্বা ম্ব ম্বা হ্ব হ্বা
Special rhotic forms		
র ru	ৰু	প্রা প্ৰা

English Consonants and Vowels			38
p	pie	pea	lower-case p
t	tie	tea	lower-case t
k	kye	key	lower-case k
b	by	bee	lower-case b
d	dye	D	lower-case d
g	gye	gye	lower-case g
m	my	me	lower-case m
n	nigh	knee	lower-case n
f	fe	fee	eng (or angma)
v	vie	V	lower-case f
θ	thigh		lower-case v
ð			theta
s	thy	thee	eth
z	sgh	sea	lower-case s
ʃ (t̬)	shy	she	lower-case z
t̬ (z̬)	shy	she	mission
l̬	le	lee	esh (or long s)
w	why	we	vision
r (r̬)	rye	{	long z (or yogh)
j (y̬)			long z̬ (or yogh)
h	high	ye	lower-case w
	Note also the following:	he	lower-case r
t̬ (t̬̬)	chime)	cheat(p)	lower-case j
dʒ̬ (dʒ̬̬)	ji(v)e)	G	lower-case h v

Hindi Vowels			41
अ a	आ ā	इ i	ई ī
उ u	ऊ ū	ও o	ঔ ū
ঢ় ঢ়া e	ঢ় ঢ়া ai	ঢ় ঢ়া au	ঢ় ঢ়া অঃ অঃ
ঢ় ঢ়া ০	ঢ় ঢ়া এঃ এঃ	ঢ় ঢ়া অঃ অঃ	ঢ় ঢ়া অঃ অঃ
ঢ় ঢ়া া	ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি
ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি
ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি
ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি
ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি
ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি
ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি
ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি	ঢ় ঢ়া ি

Vowels			39
i	heed	he	bead
ɪ	hid	bid	heat
eɪ	hayed	hay	keyed
e	head	bad	keyed
æ	had	bad	lower-case i
ɑ	hard	bard	small capital I
ɒ	hod	bird	lower-case k
ɔ	hawed	haw	lower-case b
ʊ	hood	bode	lower-case b
u	who'd	bood	lower-case t
ʌ	Hudd	bud	lower-case t
ɔ̄	herd	bird	lower-case t
aɪ	hide	high	lower-case t
əʊ	how	bowed	lower-case u
ɪə	(ahoy	Boyd	lower-case u
ɛr	here	beard	(as noted above)
aɪr	hair	bared	(as noted above)
ɔɪr	hired	hire	(as noted above)
ju	hue	Bude	(as noted above)

Indian Grammatical Tradition			42
<i>Pāṇiniyaśikṣā:</i>			
“[The <i>varṇas</i> are distinguished from each other] on the basis of <i>svara</i> “accent”, <i>kāla</i> “time, duration”, <i>sthāna</i> “point of articulation”, <i>prayatna</i> “manner”, and <i>anupradāna</i> “phonation”.”			
The oldest phonetic textbooks are the <b>Pratishakyas</b> ( <i>pratiśākhya</i> , from Sanskrit <i>prati-śākhā</i> ), describing pronunciation, intonation of Sanskrit, as well as the Sanskrit rules of sandhi (word combination), specific to individual schools or Shakhas of the Vedas. (5-4C BCE)			

43

Possible derivation of Brāhmī from the Phoenician script

Greek	A	B	T	Δ	E	Y	Z	H	Θ	I	K	Λ	M	N	Ξ	Ω	Π	Μ	Ω	Ρ	Σ	Τ	
Phoenician	𐤀	𐤁	𐤂	𐤄	𐤅	𐤇	𐤈	𐤉	𐤊	𐤋	𐤌	𐤍	𐤎	𐤏	𐤒	𐤓	𐤔	𐤕	𐤖	𐤗	𐤘	𐤙	𐤚
Aramaic	ܐ	ܒ	܂	܄	܅	܆	܈	܉	܊	܋	܌	܍	܏	ܐ	ܑ	ܒ	ܓ	ܔ	ܕ	ܖ	ܗ	ܖ	ܙ
Brahmi	ං	ඃ	඄	අ	ආ	ඇ	ඈ	ඉ	ඊ	උ	ඌ	ඍ	ඎ	ඏ	ඐ	එ	ඒ	ඓ	ඔ	ඕ	ඖ	඗	
Devanagari	अ	ब	ग	ः	ठ	व	ः	ड	ः	थ	ः	क	ः	च	ः	त्र	ः	न	ः	र	ः	त्र	
Tamil	அ	ப	க	த	ட	வ	த	ட	த	ஏ	த	ஒ	த	ஒ	த	ஏ	த	ஒ	த	ஏ	த	ஒ	
Kannada	ಅ	ಬ	ಗ	ಃ	ಠ	ವ	ಃ	ಡ	ಃ	ಥ	ಃ	ಕ	ಃ	ಚ	ಃ	ತ್ರ	ಃ	ನ	ಃ	ರ	ಃ	ತ್ರ	
IAST	a	b	g	ঃ	ঠ	v	ঃ	d	ঃ	ঠ	ঃ	k	ঃ	চ	ঃ	t্ৰ	ঃ	n	ঃ	r	ঃ	t্ৰ	
	a	ba	ga	ঃ	ঠa	va	ঃ	da?	ঃ	ঠa	ঃ	tha	ঃ	ka	ঃ	la	ঃ	ma	ঃ	na	ঃ	sa*	

Wikipedia article

46

- Note the contrast between 'resign' and 'resignation'. When the nominalizing (noun making) suffix '-ation' is added to the word, the '-g-' is pronounced. Therefore, the presence of the 'g' in the simple verb stem is not just an arbitrary quirk of spelling: the spelling reflects a **morphological** property of the word in English, which it shares with the form 'resignation'. (*Question: Why is the 'g' sometimes pronounced and sometimes not? Consider also pterodactyl and helicopter*)
- There are many other examples in which English spelling reflects morphological regularities rather than phonetic regularities. E.g., the highlighted 'c' in electric/electricity reflects the fact that the words are morphologically related, but the 'c' corresponds to a different sound in each.

44

- The **International Phonetic Alphabet** (IPA) provides a one-to-one mapping from sounds to written symbols, and can be used for any dialect of any language.
- The writing system of any language expresses speech but is not indicative of the actual phonetic values of the sounds as they occur.

47

- Ref: Peter Ladefoged's *Course in Phonetics* and *Vowels and Consonants*
- <http://phonetics.ucla.edu/>
- <https://web.uvic.ca/ling/resources/ipa/charts/IPAlab/IPAlab.htm>

45

### Phonetic Symbols vs. Orthographic Conventions

- A single sound is represented by many different spellings:

He	e	Seize	ei
Believe	ie	Seas	oe
Caesar	ae	Amoeba	ey
See	ee	Key	i
People	eo	Machine	

- A single spelling can correspond to many different sounds

dame	call	Seize	ei
dad	village	Seas	oe
father	many	Amoeba	ey

- Individual speech sounds are often represented by multiple letters

shoot	Thomas
there	deal
character	physics

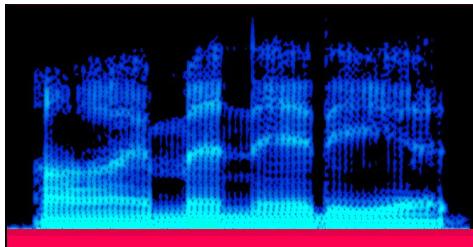
- Individual letters sometimes correspond to two speech sounds

cute [ju]	
fax [ks]	

- Letters in spelling are often not pronounced

resign	debt
psychology	pterodactyl

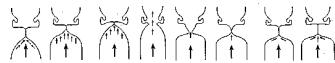
## ACOUSTIC PHONETICS



*How to read spectrograms*

1

## Vibrating glottis



**Breathy voice (or murmur):** The vocal cords are vibrating, but there is also a significant amount of air escaping through the glottis, causing turbulence.

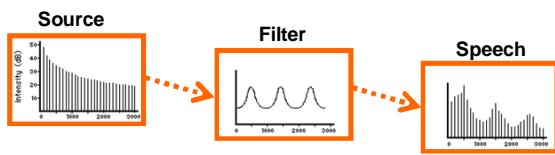
**Creaky voice:** Only the front part of the vocal folds are vibrating, yielding a very low frequency. (Try speaking at the lowest pitch you can. Then go even lower.) The very low frequency vibration of creak can be superimposed on modal voicing, producing **creaky voice**. Creak and creaky voice are often called *laryngealization* or *vocal fry*.

### Falsetto

The vocal folds are stretched tightly so that they become very thin. The resulting vibrations can have over twice the frequency that a speaker can produce using modal voicing.

4

## Source-Filter Theory of Speech Production



### Speech production can be divided into two independent parts

- Source of sound, the larynx (vocal fold vibrations includes the fundamental)
- Filters that modify the source, the vocal tract

2

- The audible frequency range for humans is 20-20000Hz. (Less than 20Hz is *infrasound* and more than 20kHz is *ultrasound*.) For dogs, 50Hz - 45kHz, cats 45Hz – 85kHz, bats can detect frequencies as high as 120kHz, dolphins as high as 200kHz. An elephant possesses the unusual ability to detect infrasound 5Hz - 10000 Hz.
- Speech sounds contain energies at all frequencies in the audible range, but most phonetic information appears to be concentrated below 8000 Hz. Telephone speech cuts off frequencies above 3500 Hz.

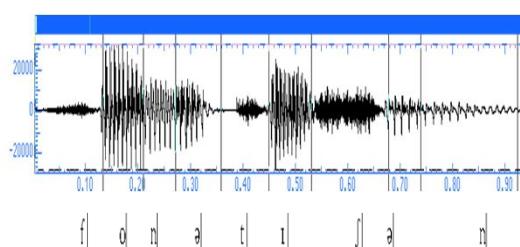
5

## Fundamental Frequency

- The typical F0 range for a male is 80-200 Hz, and for females 150-350 Hz. There is great variation in these figures
  - The F0 may be seen during voicing (when vocal folds are vibrating)
- Modal Voicing**
- The vocal folds are held together along their full length with enough tension to allow vibration: The vocal folds momentarily block airflow from the lungs.
  - The air pressure underneath the vocal folds increases.
  - The increased pressure forces the vocal folds up and apart.
  - As the pressure falls again, the vocal folds snap back together.
  - Go to first step.
- Each repetition of this cycle causes a "glottal pulse". The number of times this occurs in a second is the fundamental frequency of voice.
  - Varying the tension of the vocal folds results in different rates of vibration (and so different pitches)

3

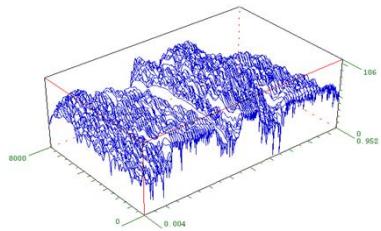
## Waveform



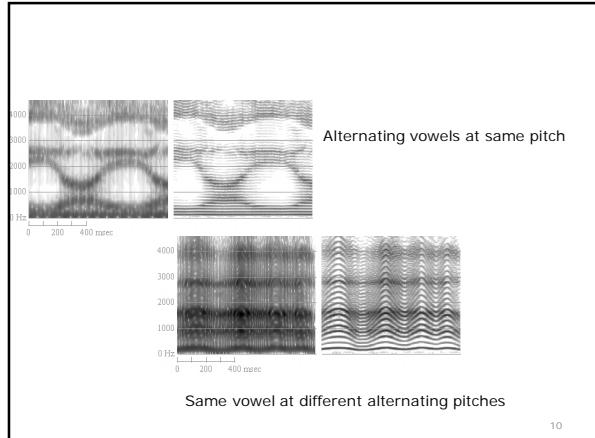
6

## Waveforms to Spectrograms

- Analyze the waveform into its component frequencies (spectra) by Fourier analysis



7



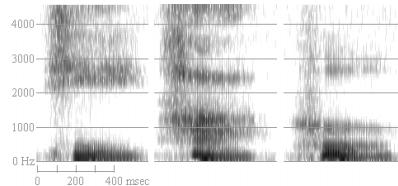
10

## Measure of the wave

- amplitude** – loudness (size of pressure differences, decibel)
- wavelength** (spatial period)
- frequency** – pitch (cycles per second, Hertz)  
Amplitude and frequency are independent of each other  
Wavelength is inversely proportional to frequency

8

## Spectrograms



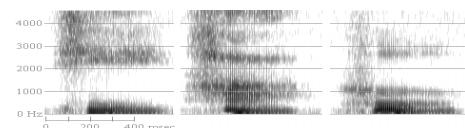
Spectrograms of "hee" "ha" and "who".

11

- Spectrogram:** Graph of the energy content of a signal expressed as function of **frequency** and **time**. Graph of a signal in which the vertical axis is frequency, the horizontal axis is time, and amplitude is shown on a grey-scale.
- Wide-band spectrogram:** A spectrogram produced using an analysis scheme which emphasises temporal changes in the signal, with short-time spectrum calculations (about 3ms)
- Narrow-band spectrogram:** A spectrogram produced using an analysis scheme which emphasises frequency changes in the signal: with long-time spectrum calculations (about 20ms).
- Each kind of spectrogram conveys different information about the speech signal.

9

## Spectrograms



Spectrograms of "hee" "ha" and "who".

**X-axis (horizontal): Time (seconds)****Y-axis (vertical): Frequency (Hertz)****Z-axis (3-d; basically, coming right toward you): Amplitude**

corresponds to signal strength, perceived as volume/loudness; higher amplitudes show up as darker marks on the spectrogram

12

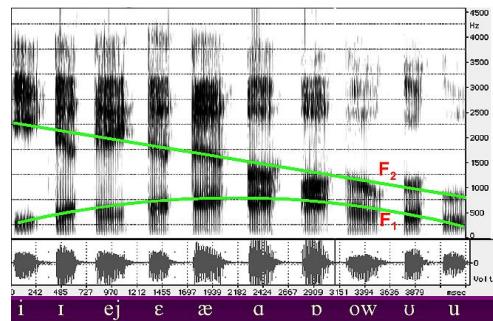
## Spectral Cues

### A. Vowels

### B. Consonants

13

## Vowels: Formants



16

## Vowels

Produced with no significant constriction of the vocal tract

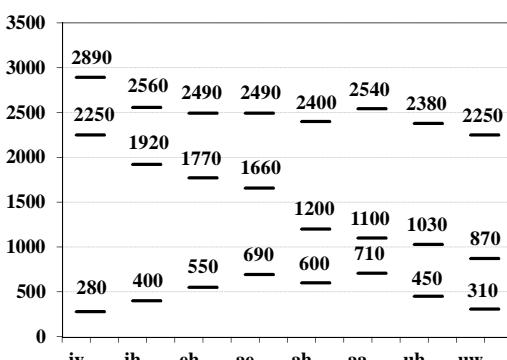
Articulations change resonances of the vocal tract

- Moving the tongue, lips and jaw changes the shape of the vocal tract

Resonances of the vocal tract are called **formants**

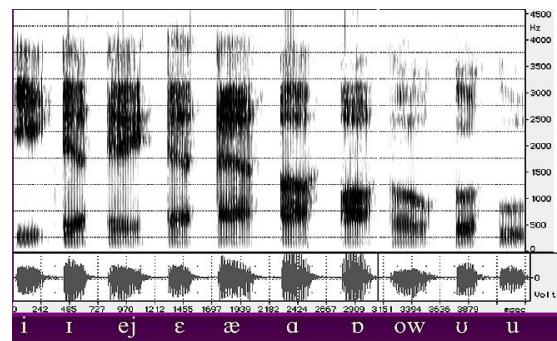
- Changing the shape of the vocal tract changes the **formant frequencies**

14



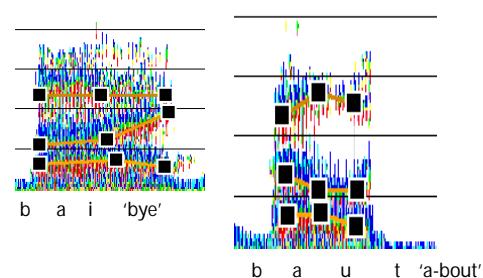
17

## Vowels: Formants



15

## Diphthongs



18

## Formant frequencies

- F1 can vary from 300 Hz to 1000 Hz. F1 correlates **inversely** to height (of the tongue). High vowels have low F1s.
- F2 can vary from 850 Hz to 2500 Hz; the F2 value correlates **inversely** with backness of the tongue. Lip rounding also causes a lower F2. Back vowels have lower F2s than front vowels.
- F3 is also important in determining the phonemic quality of a given speech sound, and the higher formants such as F4 and F5 are thought to be significant in determining voice quality.

19

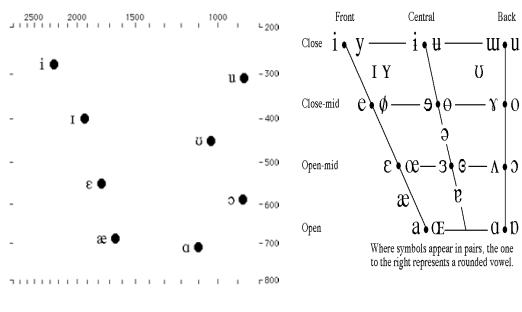
## Stops (e.g., /b/, /p/, /d/, /t/, /k/, /g/)

### Complete closure of the vocal tract at different places of articulation

- Blocks flow of air through the oral cavity
- During voiced stops, only a *voice bar* can be produced during the closure

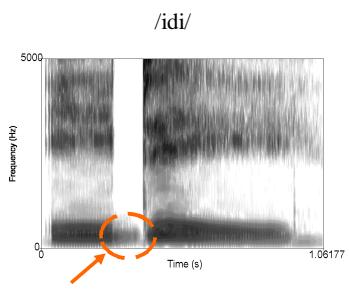
22

## IPA frequency plot and vowel chart



20

## Stops: Voice Bar



23

## Consonants

- 1. Stops**
  - Voicing
  - Aspiration
- 2. Nasals**
- 3. Fricatives**
  - Voicing
- 4. Liquids and Glides**

21

## Stops (e.g., /b/, /p/, /d/, /t/, /k/, /g/)

### Complete closure of the vocal tract

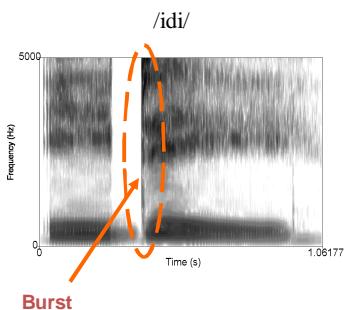
- Blocks flow of air through the oral cavity
- During voiced stops, only a *voice bar* can be produced during the closure

### Burst created at consonant release

- Created as closure is partially opened
- Frication noise created by blowing air through partial opening
- Very short duration

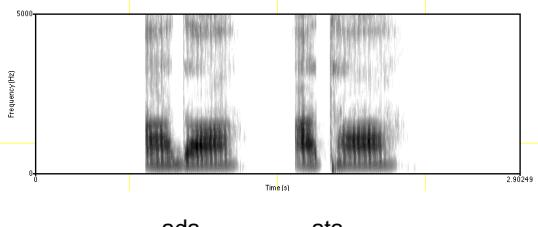
24

## Stops: Burst



25

## Stops: Voicing



28

## Stops: Voicing

### Voiced Consonants

Voicing starts less than about 30 ms after release

Voicing can occur during closure

### Unvoiced Consonants

Voicing starts more than about 50 ms after release

Voicing cannot occur during closure

26

## Voicing and Aspiration

### Voiced Consonants

- Voicing starts less than about 30 ms after release
- Voicing can occur during closure

### Unvoiced Consonants

- Voicing starts more than about 50 ms after release
- Voicing cannot occur during closure

**Duration between release (i.e. burst) and start of voicing is called Voice Onset Time (VOT).**

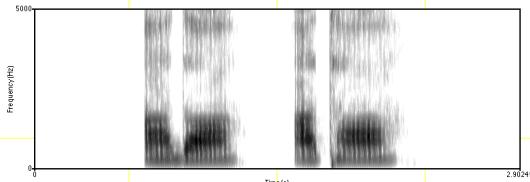
long VOT → aspiration

VOT values are set for any language and may change cross-linguistically

29

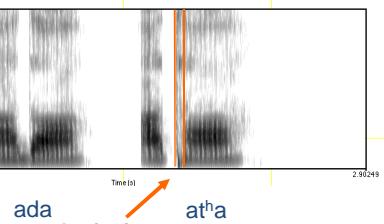
## Stops: Voicing

/ata/ vs. /ada/?



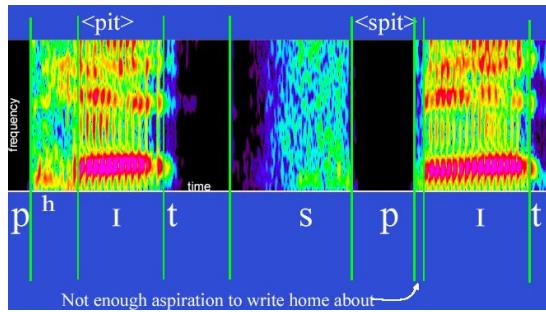
27

## Stops: Aspiration

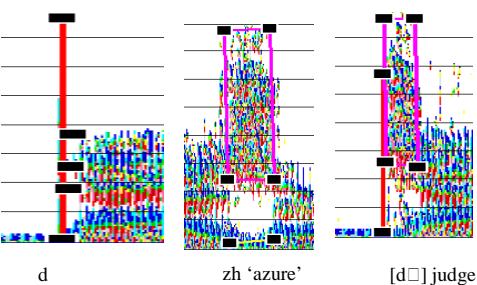


30

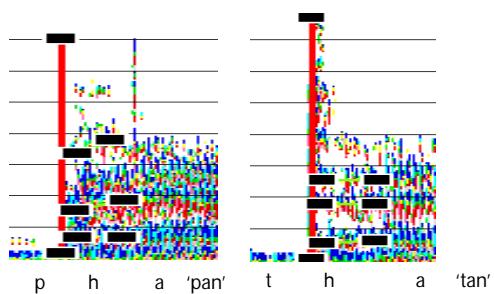
### American English Aspiration



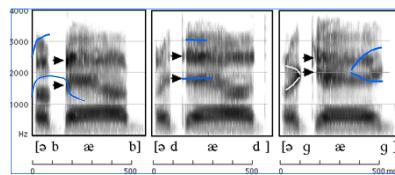
### Affricates 2



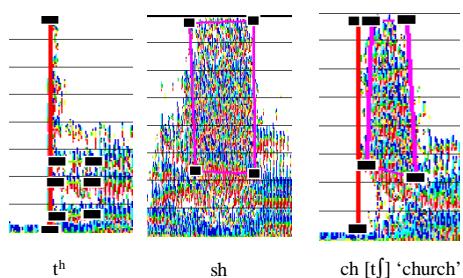
### More Aspiration



### Place and formants



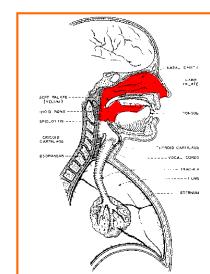
### Affricates



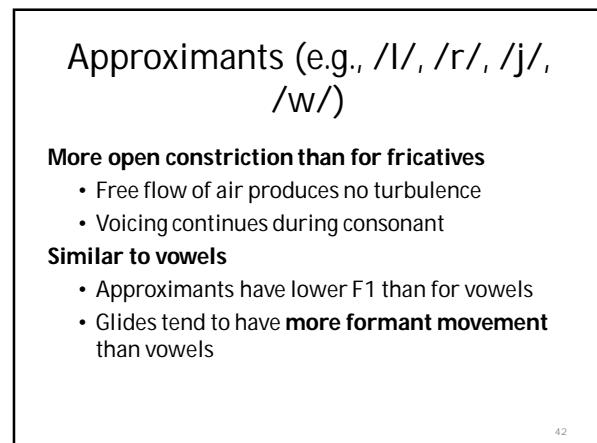
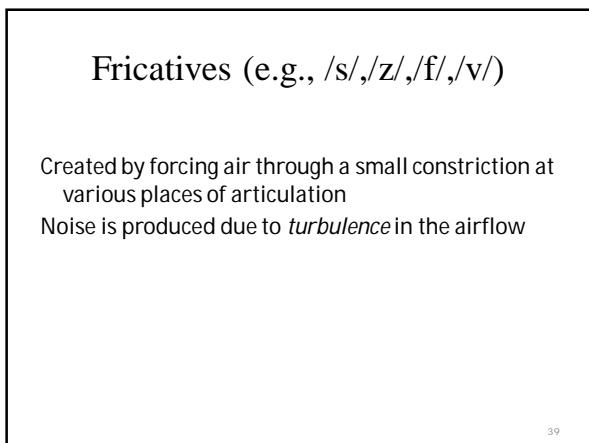
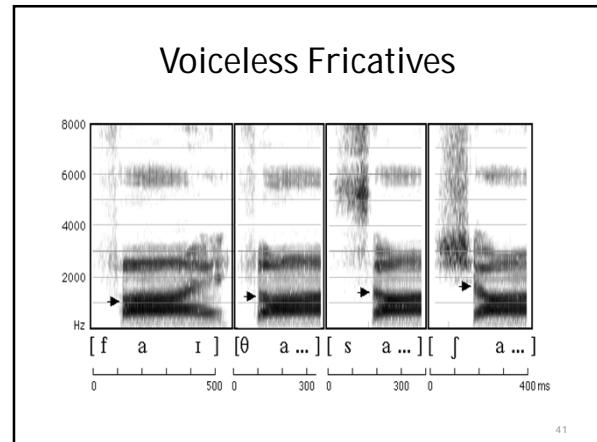
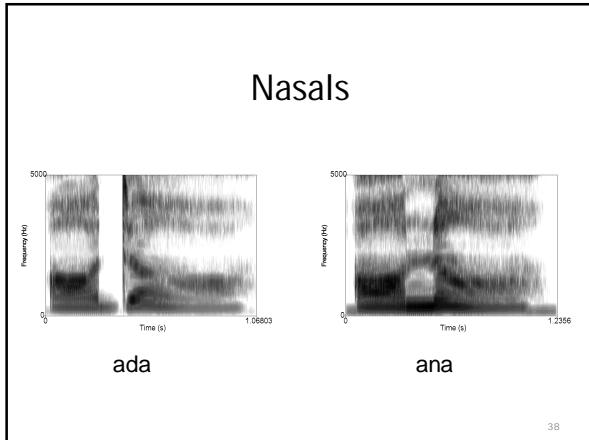
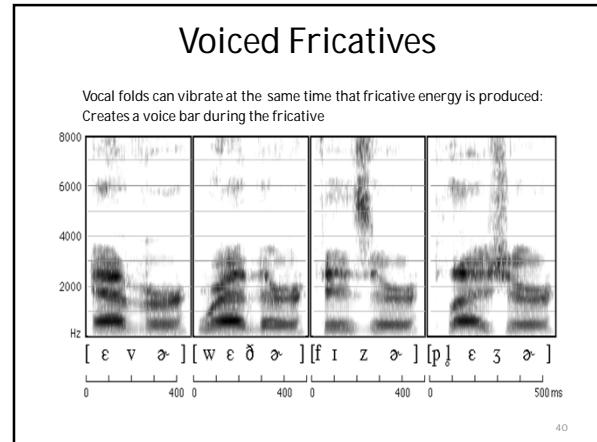
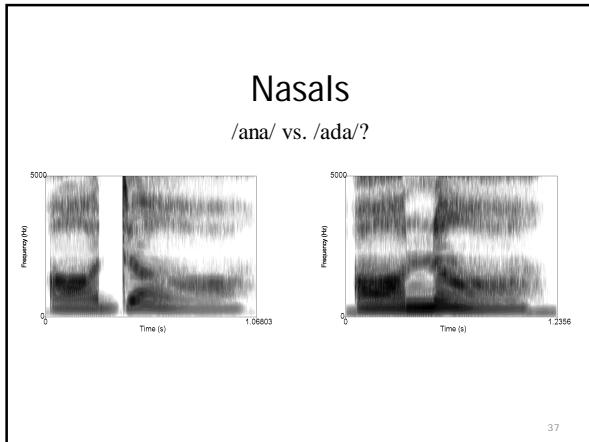
### Nasals (e.g., /m/, /n/)

Similar to voiced stops, air flows through nasal cavity

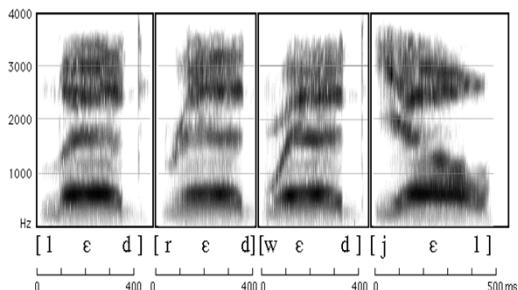
- Velum controls airflow through nasal cavity
- Airflow allows voicing to continue during closure
- Stoppage of airflow in oral tract results in decreased energy, characterized by **nasal "zeroes"** (regions of extremely low energy)



36



## Approximants



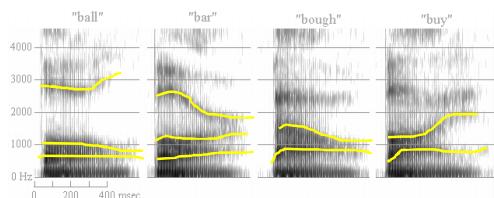
43

## Ejectives, Implosives and Clicks

- **Ejectives:** Transient (vigorous explosion) from release of articulators, followed by silence (closed glottis)
- **Implosives:** Oral closure signaled by loss of overall energy (only voicing bar) and weak energy at release; Voicing is shown by pulsing increase towards release and falling pitch from pulsing to vowel (tight vocal cords to modal position)
- **Clicks:** like plosives, characterized by a transient from release of articulators, a noise as air passes between articulators

46

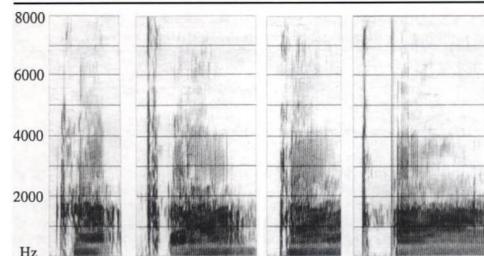
For /l/, the F2 is quite low, indicating a back tongue position—velarization of /l/ in English. The F3, on the other hand, is very high, higher than one ever sees unless the F2 is pushing it up out of the way. In "bar", the F3 comes way down, which is characteristic of [ɹ] in English.



44

## Ejectives

Velar and uvular stops and ejectives from K'ekchi



47

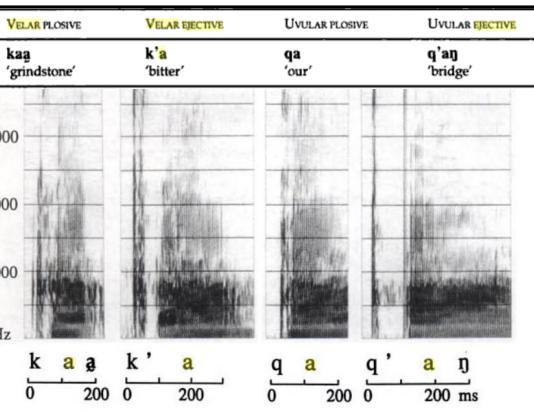
## Speech Synthesis

Speech sounds can be minimally specified in terms of a small set of parameters (variables), each of which can be described in terms of how they sound (their auditory characteristics), how they are made (physiological characteristics), or their physical (acoustic) characteristics.

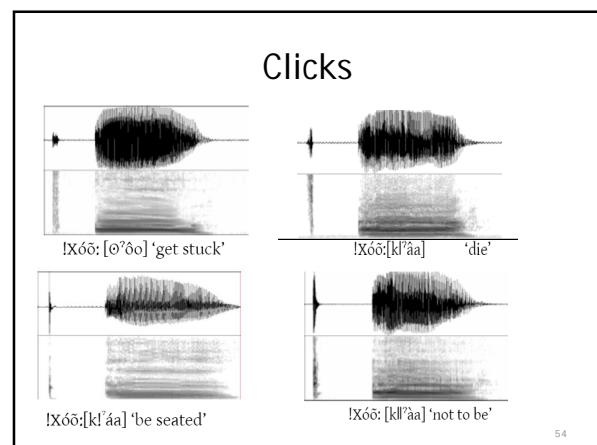
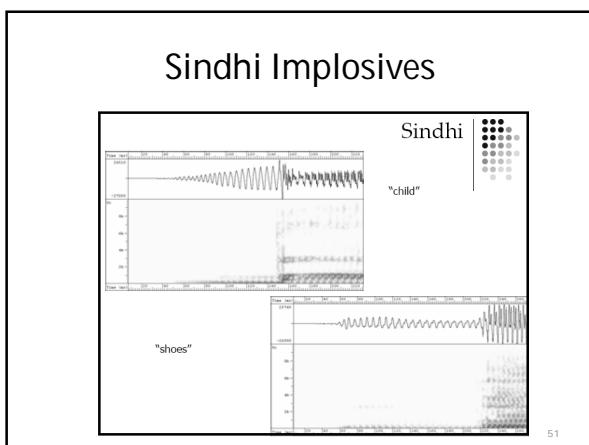
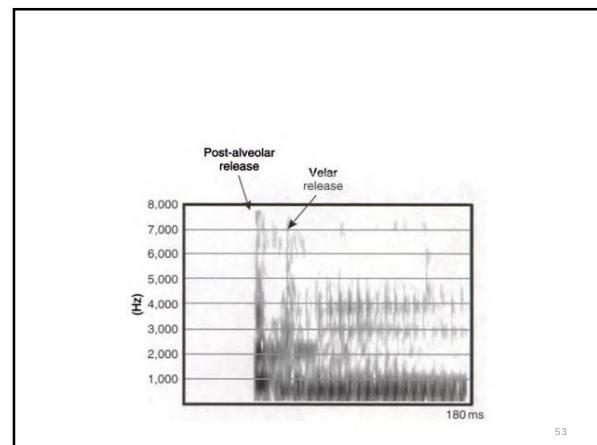
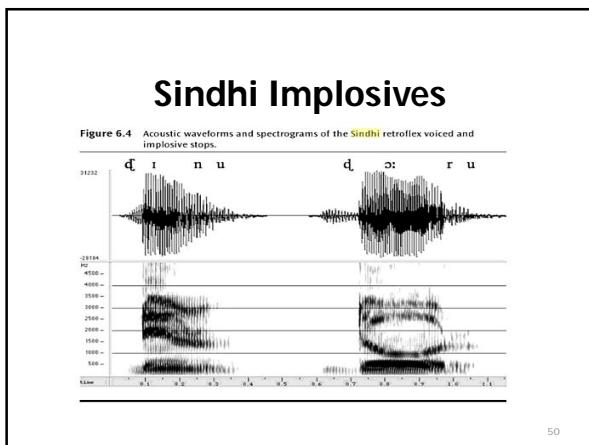
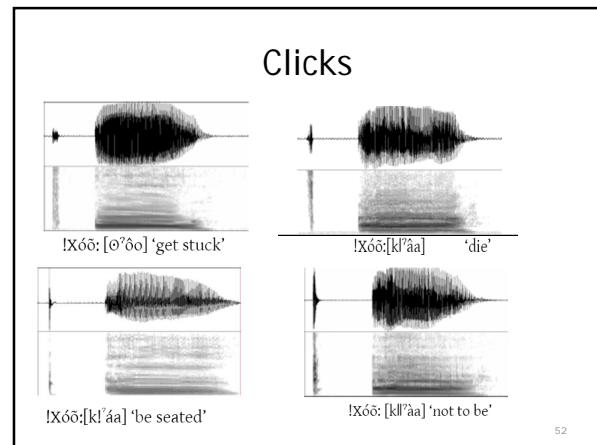
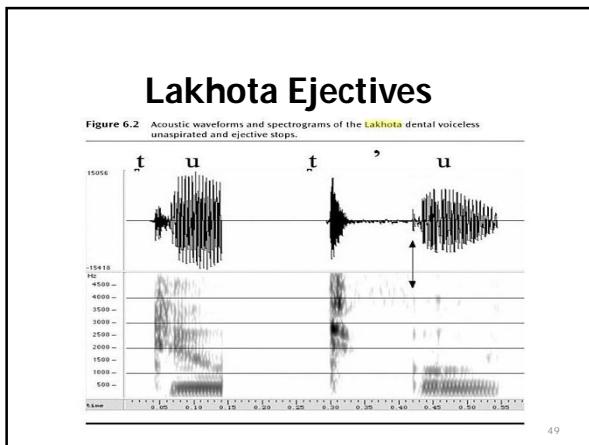
This speech was synthesized in 1971 by Peter Ladefoged on a synthesizer at UCLA. The values of the parameters were a modified version of a set provided by John Holmes.

AUDITORY	PHYSIOLOGICAL	ACOUSTIC
1 Pitch of voice	Rate of vibration of the vocal folds	Fundamental frequency
2 Overtone pitch (1)	First resonance of the vocal tract	Formant 1 frequency
3 Loudness of overtone (1)		Formant 1 amplitude
4 Overtone pitch (2)	Second resonance of the vocal tract	Formant 2 frequency
5 Loudness of overtone (2)		Formant 2 amplitude
6 Overtone pitch (3)	Third resonance of the vocal tract	Formant 3 frequency
7 Loudness of overtone (3)		Formant 3 amplitude
8 Pitch of noise	Fricative and stop bursts	Center of noise frequency
9 Loudness of noise		Amplitude of noise

The 8 links below demonstrate how speech can be built up using these parameters and additional (fixed) higher formants.



48



## Summary of Spectral Cues

Vowels: open oral tract; identified using formant structure, primarily F1 and F2

Consonants (differing degrees of constriction)

Stops: stoppage of air; voice bar; VOT

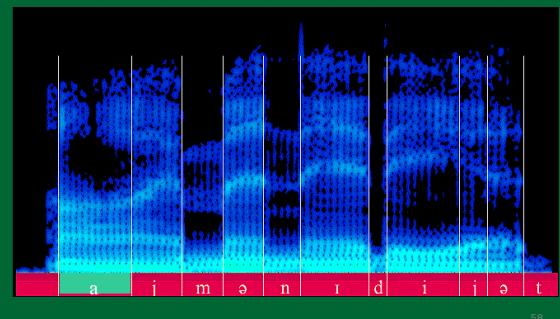
Nasals: vowel like, nasal zeroes (areas of little or low energy between formant like bands)

Fricatives: noise

Liquids and glides: vowel like, formant movement

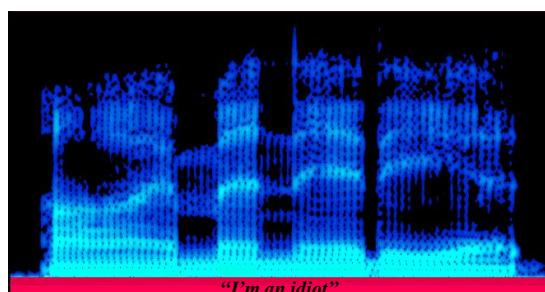
55

## Continuous Speech, Segmented



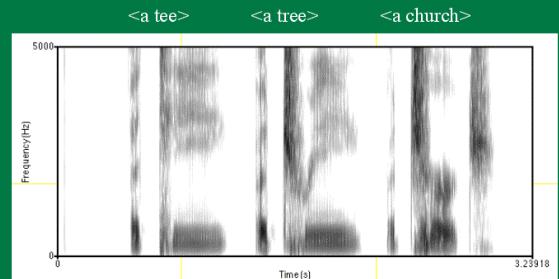
58

## Continuous Speech



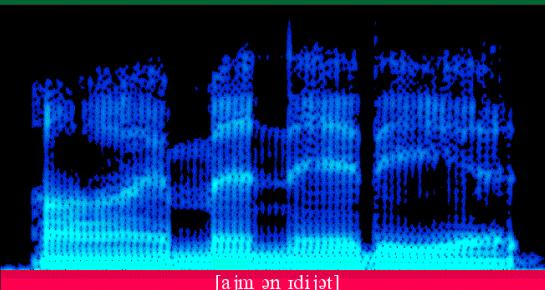
56

## Epilogue: Broad vs. Narrow Transcription



59

## Continuous Speech



57

## Web resources

- [cslu.cse.ogi.edu/tutordemos/  
SpectrogramReading/spectrogram\\_reading.html](http://cslu.cse.ogi.edu/tutordemos/SpectrogramReading/spectrogram_reading.html)
- <http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/>

60

The patterns of phones

## Phonology

1

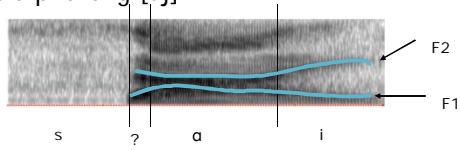
## Phonology [fənələdʒi]

- How languages deploy sounds to create words (sound-meaning pairing).
- How these sounds vary depending on their environment.
- How the sound inventories of languages are structured.
- How linguists theorize about the above.

4

## Spectrogram of *sigh*

- Note the curving formants that distinguish the diphthong [aj].



2

## Co-articulation

- Segments affect each other when in contact
- When the velar stop [k] is followed by the high front vowel [i], it moves forward in the mouth toward the alveopalatal area, in anticipation of the vowel.

5

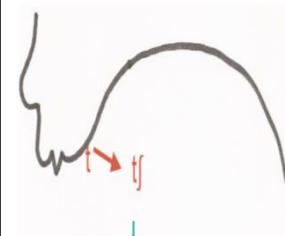
## Segments: Figments of our imagination?

- Suppose we play the vowel of *sigh*, leaving out the [s], you hear not [ai] as you might expect, but [die].
- In other words, there is information about the consonant in the following vowel. The part that is consonantal is also present in the part that represents the vowel.
- In this way, segments overlap in the stream of speech, yet we learn to hear them as purely sequential.
- We begin our study of phonology by examining some processes by which sounds are affected by their environment.

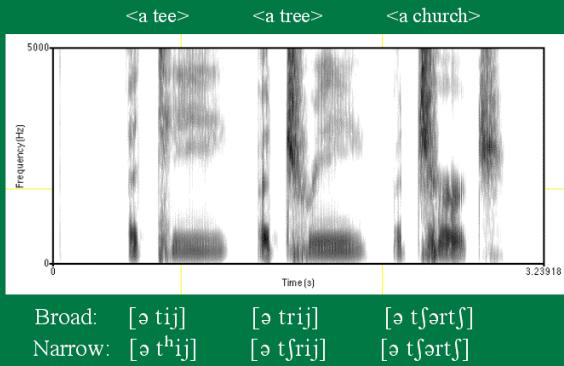
3

- the [t] in *trip, trust, intractable* opens gradually to become an alveopalatal affricate
- tʃirk
- tʃʌst
- intʃæktəbl

6



### Epilogue: Broad vs. Narrow Transcription



- If we nasalize the vowel before a non-nasal consonant in (pat, cap and sit):

pæ̃t

kāp

s̄it

It sounds odd but doesn't change the meaning

10

- In English, vowels become nasalized before a nasal coda (end of a syllable) because the velum lowers early – in anticipation of the following nasal consonant

pæt	pæ̃m	pæ̃n	pæ̃ŋ
kap	kām	kān	
sit	s̄impl	s̄in	s̄in̄g

8

- But if you do this in French or Hindi, the meaning changes:

ſa	'cat'	ſā	'field'
hɛ	'is'	hɛ̃	'are'

- Vowel nasalization is distinctive in French and Hindi but not in English.

11

- Is there any difference between the vowels in the following pairs of words?

back	bag
tat	tad
pat	pad
lap	lab
leaf	leave
plot	plop
teeth	teethe
smock	smog

9

- In English, nasalized and non-nasalized vowels don't occur in the same environments
- Nasalized vowels occur only before nasal consonants. Non-nasalized vowels never occur before nasal consonants. (Unless a speaker is being very, very careful). They are in **complementary distribution**.
- Nasalized and non-nasalized vowels are said to be **allophones** (variants) of the same phoneme

12

- In English, stop/plosive consonants are aspirated when they occur in the beginning of a syllable (onset) and are unaspirated elsewhere.

s[t]op	[tʰ]op
s[k]ool	[kʰ]ool
s[p]ort	[pʰ]ort

If you switch them and say s[tʰ]op or [t]op you may sound non-native but the meaning of the word does not change. Aspiration is not distinctive in English.

13

- And going back to the vowels in

bat	bad
tat	tad
pat	pad
lap	lab
leaf	leave

If we shorten the vowels in the second set and lengthen them in the first one, the meaning does not change. Vowel length is not contrastive in English.

16

- But in Hindi, aspiration is distinctive and aspirated and unaspirated stops are in **contrastive distribution**. They are 'phonemes' as the following minimal pairs show.

pal 'take care of'	pʰal 'knife blade'
tal 'beat'	tʰal 'plate'
ṭal 'postpone'	tʰal 'wood shop'
tʃal 'turn'	tʃʰal 'bark'
kal 'era'	kʰal 'skin'

14

- But in Danish, vowel length is phonemic

vilə	'wild'	vi:lə	'rest'
menə	'remind'	me:nə	'mean'
ləsə	'load'	lə:sə	'read'
mæsə	'mass'	mæ:sə	'mash'

- Likewise in Finnish

il	'day'	i:l	'work'
sedä	'to count'	se:da	'strong'
kul	'oyster'	ku:l	'tunnel'

What about Hindi?

17

- In Armenian too,
- |       |            |        |            |
|-------|------------|--------|------------|
| [kap] | 'bond'     | [kapʰ] | 'club'     |
| [mut] | 'entrance' | [mutʰ] | 'darkness' |
| [tak] | 'under'    | [takʰ] | 'hot'      |

15

## How do languages use sounds?

- While the human vocal tract offers us a wide variety of sounds, each language makes use of a small number of **distinctions**. It isn't the sound quality itself that does the work in language, but the set of **oppositions** among sounds.
- If two sounds can make a difference in meaning, they are separate **phonemes**.
- These distinctions are not random, but form a structured **inventory**.

18

## Complementary Distribution

- Since allophones of a single phoneme are conditioned by their environment, no two allophones of the same phoneme will occur in the same environment. In other words, their distributions will be complementary.
- Phonemic analysis involves identifying sounds that are phonetically similar enough that they constitute "suspicious pairs" and looking to see if they occur in the same (contrastive) or different (complementary) environments.
- The (phonetic) nature of the environments should explain the nature of the alternation.

19

## Finnish

[kudot]	'failures'
[kate]	'cover'
[katot]	'roofs'
[kade]	'envious'
[madon]	'of a worm'
[maton]	'of a rug'
[ratas]	'wheel'
[radon]	'of a track'

Can you tell if [t] and [d] separate phonemes?

22

## Context Sensitive Rules

- Rule:  
o → o / \_ Nasal  
(the segment or segments that undergo the change, the kind of change and the environment which induces the change)
- Picking a base form
- Making the rule as general as possible

20

## Swahili

[ŋɔma]	'drum'	[watoto]	'children'
[bɔma]	'fort'	[ndoto]	'dream'
[ŋɔmbe]	'cattle'	[mboga]	'vegetable'
[bɔmba]	'pipe'	[ndogo]	'little'
[ɔmba]	'pray'	[dʒogo]	'rooster'
[ɔna]	'see'	[ʃoka]	'axe'

Are [ɔ] and [o] separate phonemes?

23

## Writing Rules

- In English, [e] becomes [ɛ] before a nasal consonant coda. And [æ] becomes [æ̃], [a] becomes [ã], [o] becomes [õ], and so on.
- We could write a rule for each of these vowels:  
[e] -> [ɛ] / \_ Nasal  
[æ] -> [æ̃] / \_ Nasal  
[o] -> [õ] / \_ Nasal  
and so on down the line
- But if we did, we'd be missing an important generalization - that ALL vowels become nasalized in this position:  
V -> V / \_ Nasal
- We try to make our rules as general as possible

21

## Japanese

[č], [t], [ts]

tatami	tomodači	uči
tegami	totemo	otoko
čiči	tsukue	tetsudau
šita	ato	matsu
natsu	tsutsumu	čizu
kata	tatemono	te

24

### Mokilese

p̄isan	'full of leaves'	
dup̄ukda	'bought'	Micronesian
p̄uko	'basket'	Language
k̄isa	'we two'	Region: Mokil
sup̄wo	'firewood'	Atoll
kamwɔk̄jti	'to move'	
uduk	'flesh'	
kaskas	'to throw'	
poki	'to strike something'	
pil	'water'	
apid	'outrigger support'	
lujuk	'to tackle'	

25

br̄en	'endanger'	gat	'cat'
bako	'cow'	l̄ung	'long'
ūmbro	'shadow'	saliþo	'saliva'
kr̄ambo	'room'	noþi	'husband'
dilys	'Monday'	aþe	'to have'
dūŋko	'until'	siþaw	'horse'
duso	'sweet'	byðet	'gut'
taldepān	'leftover bread'	eþaðo	'hoe'
p̄unde	'to lay eggs'	biyar	'mosquito'
dudze	'twelve'	riyut	'he laughed'
guteþa	'flow'	agro	'sour'
ēngwan	'this year'	juyet	'he played'
puðe	'to be able'		

28

### Greek

[χ] voiceless velar fricative  
 [κ] voiceless velar stop  
 [ç] voiceless palatal fricative  
 [c] voiceless palatal stop

[kano]	'do'	[kori]	'daughter'
[xano]	'lose'	[xori]	'dances'
[çino]	'pour'	[xrima]	'money'
[cino]	'move'	[krima]	'shame'
[kali]	'charms'	[xufta]	'handful'
[xali]	'plaint'	[kufeta]	'bonbons'
[çeli]	'eel'	[oci]	'no'
[ceri]	'candle'	[ceri]	'hand'

26

### Osage (Great Plains)

1. [dabri] 'three'
2. [datspe] 'to eat'
3. [dak?e] 'to dig'
4. [dal?] 'good'
5. [daſtu] 'to bite'
6. [aðikhāzā] 'he lay down'
7. [ts?eðe] 'he killed it'
8. [ðeze] 'tongue'
9. [ðie] 'you'
10. [ðuða] 'to wash'

Are [d] and [ð] separate phonemes?

29

### Gascon (Southwestern France)

What is the phonemic status of the voiced stops [b,d,g] and voiced fricatives [β, ð, γ]? What's the most general rule you can write for this alternation?

27

### Zulu (South Africa)

1. bōna	'see'	13. iboni	'grasshopper'
2. bōpha	'bind'	14. umondli	'guardian'
3. mōsa	'despoil'	15. umosi	'one who roasts'
4. umōna	'jealousy'	16. inoni	'fat'
5. imōtɔ	'car'	17. udoli	'doll'
6. iqɔlɔ	'small of back'	18. umxoxi	'story-teller'
7. ixɔxɔ	'frog'	19. imomfu	'jersey cow'
8. isicɔɔɔ	'head ring'	20. lolu	'this'
9. isithōmbe	'picture'	21. isitofu	'stove'
10. Indōdāna	'son'	22. nomuthi	'and the tree'
11. umfɔkazi	'strange man'	23. udodile	'you acted like a man'
12. ibokisi	'box'		

What is the distribution of [o] and [ɔ]?

(/b/ is a voiced bilabial implosive stop. /c q x/ are clicks.

30

## English

p[r]ay	b[r]ay	p[ɿ]ip	b[l]ip
t[r]ay	d[r]ay	(*tɿ)	(*dl)
c[r]ay	g[r]ay	c[ɿ]ass	g[l]ass
sh[r]ed			
		f[ɿ]og	
		s[l]og	

31

## Points to note

- Not all *phones* (speech sounds) of a language are in contrast. Some phones are predictable variants phonetically similar systematic occurrence is rule-governed, predictable
- Distribution of phones: contrastive or complementary  
Contrastive → Phonemes  
Complementary → allophones
- What is a phoneme and what are allophones is language specific and arbitrary; but the processes themselves are repeated in language after language.

34

## English [l]

life	poll
leap	pill
lose	foal
elope	full
delight	help
sleep	bulk

eligible, legal, level, grumble vs. grumbling

32

## What is a phoneme?

Phoneme: Speaker's **internalized mental representation** of a single speech sound

Reality of phonemes:

(unit of grammatical organization)

- speech errors target segments i.e., phonemes
- perception (of infants and native speakers) is of phonemes
- language games

35

- What about Hindi nasals  
ङ अ ण न म or anusvar
- Or  
শ vs ষ

33

## phonological processes and changes target segments

### insertion or epenthesis

something >some[p]thing  
warmth >warm[p]th  
length >leng[k]th  
tenth >ten[t]th

anros > andros (androgynous)  
amrotos > ambrotos (ambrosia)  
ganra>gandra>gander  
simle>simble>simple  
æmtig> æmptig>empty

k[u]lub  
t[i]ren  
s[l]port

36

**deletion**

[p]psychology, [p]salm, [p]seudo, [k]night, [k]not, [k]nit  
 kamār > kamroN, nazar > nazroN, gagan > gagnoN, pakāD > pakDa

**metathesis or transposition of segments**

bridd > bird	waps > wasp
thridd > third	ask > aks
frist > first	prescribe > perscribe
hros > horse	

- Latin *parabola* > Spanish *palabra* 'word'
- Latin *miraculum* > Spanish *milagro* 'miracle'

37

**Major features (binary values)****Class Features**

- [+/- consonantal] (oral constriction)
- [+/- sonorant] (resonance vs turbulence)
- [+/- syllabic] (syllable nucleus)

**Manner Features**

- [+/- continuant] (no oral obstruction)
- [+/- lateral] (center of tongue touches roof)
- [+/- nasal] (velum lowers)
- [+/- strident] (high energy white noise)

40

**dissimilation**

Latin – Spanish – Italian  
 anma > alma 'soul' > anima > âme  
 arbor > arbol (Sp) > alboro (It), > arbre (Fr)  
 Latin *rārus* "rare" → Italian *rado*.

**reanalysis**

a numble pie > an umble pie > humble pie  
 a narang > an orange, a napron > an apron

**phonological split**

n > n and ng (singen-sing)

38

	plosives	fricatives	nasals	liquids	glides/vowels
[consonantal]	+	+	+	+	*
[sonorant]	*	*	+	+	+
[continuant]	*	+	*	+	+

41

**Features**

- Individual phones are described as a sum of features (phonetic), and these **distinctive features** are the building blocks of phonology
- Features describe **natural classes**
- Phonological rules operate on natural classes or are triggered by specific features of the segments or the context

39

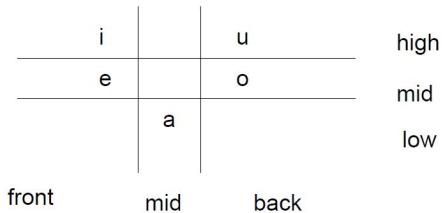
**Some others**

- Laryngeal Features**
  - [+/- voice] (periodic vibration of vocal cords)
  - [+/- aspirated] (spread glottis)
  - [+/- glottalic] (constricted glottis)
- Place Features**
  - [+/- round]
  - [+/- high]
  - [+/- low]
  - [+/- back]
  - [+/- tense] ('advanced tongue root')
  - [+/- anterior] (tip of tongue in front of alveolar ridge)
  - [+/- distributed] (tongue extended in mouth)

42

## Vowels

### a common vowel system



What's odd about this vowel system?



46

### using features

i		u	+ high -low
e		o	- high -low
	a		- high +low
- back		+back	

## Above the segment

- Segments are grouped into another unit of grammatical organization called the **syllable**
- Usually built around a **vowel** (and sometimes liquids and nasals) with associated consonants
- Intuitive and reliable demarcation of words into syllables (not into random clusters)

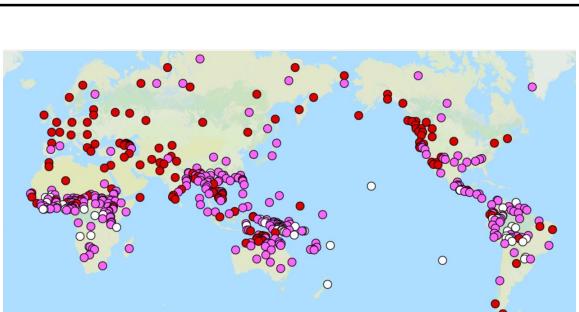
floccinauccinihilipilification  
Floc-ci-nau-ci-ni-hi-li-pi-li-fi-ca-tion

47

What's odd about this consonant system?

p	t	k	p	t	p <sup>h</sup>
b			b	d	b <sup>h</sup>
f	s	x	f	s	
v	z	y	b	z	
m	n	ŋ	m	n	

45



THE WORLD ATLAS  
OF LANGUAGE STRUCTURES  
ONLINE

48

## Phonotactics

- Permissible syllable structures, changes for each language
- |                     |                 |               |
|---------------------|-----------------|---------------|
| <u>Russian</u>      | <u>Hawaiian</u> | <u>Hebrew</u> |
| fprok 'value, good' | CV              | CV(C)         |
| psa 'dog's'         | V               | CCV           |
| fslux 'aloud'       |                 | CCVC          |
| pt̪itsə 'bird'      |                 | CVCC          |
- What are the kinds of onsets that are possible in English?

49

- /sɪ.'læ.bi.fl'.keɪ.ʃən/

52

- Initial consonant clusters in English

Labial+sonorant	Coronal+sonorant	Velar+sonorant
[p]lease	[tʃ] -	[k]ean
[pr]oud	[tr]ade	[kr]eam
[pw] -	[tw]in	[kw]ick
[py]ure	[tʃ]une	[kj]ute
[spl]at	[stʃ] -	[sk]erosis
[spw] -	[str]eet	[skw]eak
[spj]ew	[stw] -	[skj]ewer

- s-[pt,k]-[l,r,w,y]
- 24 possible 2 and 3 consonant clusters

50

## Reconsidering the rules

- Vowel length increase  
o-bey, re-do, re-gard, o-gre)  
(bat:bad, boat:bode, leak: league:ea-ger)
- Aspiration, unreleased stops, unaspirated stops  
[pʰ]it, s[p]it, si[p']
- [l] alternation  
pill, feel, help, bulk, sell, full vs.  
leaf, leap, lose, slope, delight  
le.vel, le.gal, la.pel

53

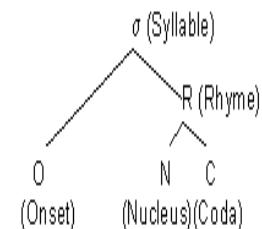
## How to syllabify?

- Find the vowels, nuclei usually [sai.kɔ.lə.dʒɪst]
- Sonority Sequencing Principle:  
The sonority of a syllable increases from the beginning of the syllable to the peak, and decreases from the peak to the end of the syllable.  
stops < fricatives < nasals < liquids < glides < vowels
- Maximal Onset Principle:  
Then attach consonants to the onset  
Remaining consonants into the coda  
Accidental gaps, *klof, flin, tron, plimt*  
Systematic gaps, *bzin, fnost, gban*

51

## Structure of the syllable

Vowels occupy the nucleus, also called peak; consonants occupy the onset and coda positions. Together the Nucleus and coda constitute the rhyme/rime. Poetic rhyme pays attention to the NC sequence – fun, sun, run etc.



54

## Stress: stress-timed languages and poetry

- Syllables are grouped into constituents called feet (sg. Foot) or metrical feet  
~~ iamb ~~ trochee, ~~ spondee, ~~~ dactyl ~~~ amphibrach ~~~ anapest

Shelley's *Stanzas Written in Dejection, Near Naples*

Alas! I have nor hope nor health,  
Nor peace within nor calm around  
Nor that content, surpassing wealth  
The sage in meditation found,  
And walk'd with inward glory crown'd;  
Blake's "Introduction" to *Songs of Innocence*:  
PIPING | down the | valleys | wild,  
Piping | songs of | pleasant | glee,  
On a | cloud I | saw a | child,  
And he | laughing | said to | me:

55

## Stress: Bounded feet

Alternating Stress languages, scanning L to R for feet

Maranungku (Australia)

(x .) (x.) (x)

(σ σ)(σ σ) (σ)

**laŋka ra te ti** *prawn*

Chemehuevi (California)

(. x) (. x) (x)

(σ σ)(σ σ) (σ)

**un i nupi ru** *make*

(x .) (x .) (x .)

(σ σ)(σ σ) (σ σ)

**wele pe le manta** *kind of duck*

(. x) (. x) (. x)

(σ σ)(σ σ) (σ σ)

**ha?i sutu iv i** *s/he liked*

Head: R

Head: L

58

- Byron's *Destruction of Sennacherib*

The Assyrian came down | like the wolf | on the fold,  
And his cohorts were gleam | ing in pur | ple and gold;  
And the sheen | of their spears | was like stars | on the sea,  
When the blue | wave rolls night | ly on deep | Galilee.

- Thomas Hardy's, *The Voice*

Woman much | missed, how you | call to me, | call to me,  
Saying that now |you are not |as you were  
When you had |changed from the |one who was |all to me,  
But as at | first, when our | day was fair.

56

### Parameters:

(1) Build feet L to R

Stress R

(2) Stress L

Even σ stressed

(3) Odd σ stressed

59

## Stress: Unbounded feet

### French

ou  
ami  
emporter  
amicalement  
hospitalisation

### Hungarian

kévert 'mixed'  
kéverget 'keep stirring'  
kéveretlen 'unstirred'

57

## Bounded feet

### Scanning R to L

Warao (Venezuela)

(x .) (x.)

(σ σ)(σ σ)

**wara wara** *kind of fish*

Weri (New Guinea)

(. x) (. x)

(σ σ)(σ σ)

**ul u amit** *mist*

(x) (x .) (x .)

(σ) (σ σ) (σ σ)

**ji wara na e**

*he found it*

(x) (. x) (. x)

(σ)(σ σ) (σ σ)

**a kune tepal**

*times*

60

## R-L Odd no. of syllables

### Maranungku (Australia)

(x .) (x .) (x)  
 $(\sigma \sigma)(\sigma \sigma)(\sigma)$   
**lan**ka **ra** te **ti** *prawn*

(x) (x .) (x .)      (x) (. x) (. x)  
 $(\sigma)(\sigma \sigma)(\sigma \sigma)$   
**ji** **wara** **na** **e**      **a** **kune** **tepal**  
*he found it*            *times*

### Chemehuevi (California)

(. x) (. x) (x)  
 $(\sigma \sigma)(\sigma \sigma)(\sigma)$   
**un** **i** **nupi** **ru** *make*

61

## Syllable weight - Latin

Stress the antepenultimate syllable

*clepsydra*, -ae. nf. 'water clock'

*formula*, -ae. nf., nice shape, beauty, (math.) formula, equation, identity.

*femina*, -ae. nf., woman

except if the penultimate syllable is heavy:

■ Long vowel VV: *puellārum* (girl),

■ Diphthong: *filiābus* (daughter)

■ Consonant in coda, *spelunca*, -ae. nf., cave.

64

### Parameters:

(1) Build feet R to L

Stress R

(2) Stress L

1+even σ stressed

(3) 1+odd σ stressed

62

## Syllable based processes: Tone

### Tone as a Morpheme: Mende Noun Classes

	1 σ	2σ	3 σ
H	kó 'war'	pélé 'house'	háwámá 'waistline'
L	kpà 'debt'	bèlè 'trousers'	kpákàlì 'tripod chair'
HL	mbú 'owl'	ngilà 'dog'	félàmà 'junction'
LH	mbá 'rice'	nàvó 'money'	ndávúlá 'sling'
LHL	mbā 'companion'	njähà 'woman'	níkílì 'groundnut'

65

## Word level stress

Word level ( x . . )

Foot level ( . x ) ( . x ) ( . x )

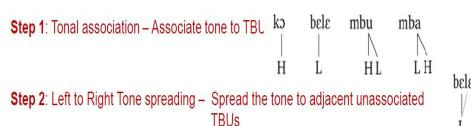
Feet (σ σ)(σ σ)(σ σ)

Syllables ha?i sutu iv i s/he liked

Output: **ha?**ísutuiví

Primary and Secondary stresses

63



66

### Mende Suffixes

The following data illustrate three **suffixes** in Mende (Sierra Leone). Unsuffixed stems are given on the left. Tones are indicated as follows:

- á = high
- à = low
- â = falling
- ã = rising
- 'á = high preceded by downstep

67

### 'on'

1. kó	'war'	kó'má	H-H
2. pélé	'house'	pélémá	H-H
3. mbú	'owl'	mbúmá	HL
4. ngílá	'dog'	ngílámá	HL-L
5. mbā	'rice'	mbámá	HL
6. bélè	'trousers'	bélémá	L
7. nyáhá	'woman'	nyáhámá	LHL
8. návó	'money'	návómá	LH-H
9. fändé	'cotton'	fändémá	LH-H

fändé      fändé -má

68

### indef. plural

kó ngá:	-ngá	kó'i
pélengá:		pélé'i
mbúngá:	HL-L Two Ls merge	mbú'i
ngílángá:		ngí'lé'i
mbängá:	LH+L	mbé'i
bélengá:		bélé'i
nyáhangá:	LHL + L Two L's merge	nyáhé'i
návóngá:		návói
fändengá:		fändé'i

### def. singular

69

## MORPHOLOGY: STUDY OF THE STRUCTURE OF 'WORDS'

### Word

- What is a word anyway?
- What do you know when you know a word?
- Do you have to learn every single word in the language?
- Are there pauses between words?  
a door/adore; a maze/amaze.
- Is each word a single chunk of meaning?  
"writer" and "a person who writes"  
"lamb" and "young sheep".

### Best definition

- A word is a "minimal free form", words can't be interrupted:  
the writers    the good writers    \*the writ good ers
- Words can move as units:  
John called up his mother. John telephoned his mother.  
John called his mother up. \*John teled his mother phone.
- The relation of words to surrounding elements is not fixed—at least not entirely  
the happiest dog    the dog that's happiest \*the est happy dog
- Words can occur in isolation :  
Who are they?    Painters.  
What do they do?    Paint.  
Who paints?    \*-ers
- What about highly synthetic languages?

### Parts of Speech

- 4 □ What parts of speech are the following?  
gleam  
gleaming  
luminous  
glow
- On what basis are you making this judgement?
- Knowledge of lexical/syntactic category (*part of speech*)  
The sea {gleamed/\*luminoused/ \*gleaminged/glowed).  
I saw a {gleam/\*luminous/gleaming/ glow} on the horizon).  
There was a {\*gleam/luminous/ gleaming/\*glow} haze on the horizon.

### Words and Morphemes

- 5 □ Do you have to memorize every single word/word form?  
paint, painter, painters  
swim, swimmer, swimmers  
blick, blicker, blickers (blick, to move arms up and down together)
- Morphemes: minimal linguistic signs, the atoms of meaning
  - An (arbitrary) *sound-meaning* pair that cannot be broken up into smaller segments that have meaning
  - both the sound and the meaning are crucial to the identity of a morpheme
  - ≠ single syllable, Hippopotamus, happy, river, apocalypse
  - ≠ words, words may be monomorphemic (unanalyzable) or (analyzable) polymorphemic

### The Mental Dictionary or Lexicon

- 6 **Definition 1: The Mental Dictionary contains all the words you've heard (or read) before.**
- Problem: It fails to account for our ability to easily understand (and create) words that we have never heard before.
  - squishability
  - tealish
  - unimpeach
  - googolbyte
- Figuring out the plural form (for nouns) or the past tense (of verbs) of unheard of words.
  - dax (plural: daxes; past tense: daxed)
  - bingle (plural: bingles; past tense: binged)
  - riffett (plural: riffets; past tense: riffeted)

## The Mental Dictionary or Lexicon contd.

7

**Definition 2:** The Mental Dictionary contains all the words you can know (previously and not yet heard).

- Problem: People can (potentially) know more words than they could likely ever store in their heads.
- Example 1
- Verbs in **Spanish**: 6 person/number forms per verb, for each of 7 tenses = 42 forms per verb
  - **entrar** 'to enter'
- 1sg: entro (present tense)  
2sg: entras  
3sg: entra  
1pl: entramos  
2pl: entráis  
3pl: entran

## The Mental Dictionary or Lexicon contd.

10

**Definition 3:** The Mental Dictionary contains every meaningful word-piece or list of morphemes

dog  
cat  
velociraptor  
un-  
-able  
-ity  
-ish

- Problem: What are legitimate combinations and what aren't?

\*run-ity  
\*funn-er  
\*disantimentestablishismarian

## The Mental Dictionary or Lexicon contd.

8

Example 2

- Verbs in **Nimboran** (Papua New Guinea)
  - 4 tenses: future, present, recent past, distant past
  - 5 persons: 1st, 2nd, 3masc, 3neutral, Inclusive
  - 3 numbers: singular, dual, plural
  - 15 locations: above, below, there, far away etc.
  - 2 aspects: repeated, not repeated
  - subject agreement, object agreement
- About 4000+ forms per verb
 

ngeguomansedam  
ngeduo - man - se - d - am  
draw-incl.dl.subj-7loc-fut-inc  
'You and I will draw from here to there'

11

**Definition 4:** The Mental Dictionary contains morphemes and instructions for how to combine them

- The Unpredictability Principle:

Only list unpredictable information in the mental lexicon. Clausal meaning is predictable:

Ram likes Sita.  
The monsoon is here.

- Compositional meaning: The meaning of a phrase is a composite of the meanings of the words in that phrase.
- Problem: Although clauses have compositional meaning, some larger-than-a-word units don't.

## Nimboran is crazy

9

- Cardinality: How many words does English have?
  - An Infinity!
    - missile: 'ICBM'
    - anti-tank-missile: 'missile targetting tanks'
    - anti-aircraft-missile: 'missile targetting aircraft'
    - anti-missile-missile: 'missile targetting ICBMs'
  - anti-anti-missile-missile-missile
  - anti-anti-anti-missile-missile-missile-missile:  
'missile targetting anti-anti-missile-missile-missiles'
- And so on...

## Larger than a Word Entries

12

- Idioms: The meaning is not predictable from the meanings of the constituent words. Idioms have non-compositional meaning, and must be listed in the mental lexicon.
  - He let the cat out of the bag.
  - Break a leg!
  - He kicked the bucket.
- Compounds: Meaning of compounds is related to their parts
  - baby powder corn syrup TV antenna foot- rest
  - storybook pocketbook wireless local loop
  - baby powder = powder to put on babies
  - corn syrup = syrup made from corn
  - TV antenna = antenna attached to a TV
  - storybook = book filled with stories

13

Are compounds compositional?

- story book = book filled with stories
- cheque book = book filled with cheques
- school book = book used in school (\* book filled with schools)
- bank book = book issued by a bank (\* book filled with banks)
- baby book = book about babies (\* book filled w/ babies)

Can you predict the meaning of new compounds?

- |             |                |
|-------------|----------------|
| fish powder | screen glare   |
| cat print   | computer model |

16

Morphology is the study of meaningful units (morphemes) in language and how they combine to form words. Complex words can be decomposed into smaller (meaningful) units.

And a morpheme is a minimal linguistic unit that carries a constant, indivisible (non compositional) meaning or a grammatical function.

Morphemes with **Grammatical Function** either signal relationships between words in a sentence or convey meanings so vital and basic to communication that they get used repeatedly.

14

## Smaller-than-a-word entries

15

 Some words have compositional meaning:

- |       |        |      |        |
|-------|--------|------|--------|
| mynah | mynahs | walk | walked |
| snake | snakes | hunt | hunted |

Meaning of "mynahs" = meaning of "mynah" + meaning of "- s".

Meaning of "walked" = meaning of "walk" + meaning of "- ed".

 Mental Dictionary: Summary must contain -

1. words- sized units (cat, open, etc.)
  2. bigger- than- a- word units (idioms, compounds)
  3. smaller- than- a- word units (affixes)
  4. Rules for combining 1 and 3
- etc.

17

 Signaling Relationships between words

The box is on the table.

John is taller than you.

Patricia's book

 Signaling Vital or Basic Meanings

- |        |         |         |
|--------|---------|---------|
| dog    | ~       | dogs    |
| danced | ,       | dancing |
| I run, | he runs |         |

16

## What Goes into an Entry?

17

At least:

- The sounds that make up the entry (*phonological string*)
- What the entry means (ex: cat means CAT), *semantic features*  
(Note: This sound-meaning pairing is an arbitrary one)
- Grammatical features*, gender, number...
- Irregularities (ex: goose-\*gooses-geese; go-went)
- Lexical category: noun, verb, etc.
- Subcategorization:  
intransitive, transitive, ditransitive, optionally transitive...  
selection restrictions (ex: must have an animate object--\*  
the table lied to me)

18

Problems: **1. Some things that look like morphemes aren't:**

- relock = re- + lock "to lock again"
- reread = re- + read "to read again"
- retype = re- + type "to type again"
- rehearse = re- + hearse "to hearse again"
- reply = re- + ply "to ply again"
- retort = re- + tort "to tort again"

What do "hearse", "ply" and "tort" mean?

- rehearse < Middle French rehercier: hercier: to strike, to plow re-hercier: to plow through something over and over again. (We borrow the word, not the morphemic structure.)

19

□ **2. Different morphemes can look the same**

There are three "s" morphemes in English:

Plural -s: cat > cats

3rd Singular -s: open > opens

Possessive -s: Joe > Joe's

1. Different meanings (plural, 3rd sing., possessive)
2. Attach to different types of words (nouns, verbs)

□ **The Morpheme Test:**

- (1) Does the meaning of the word as a whole, exactly equal the combination of the meanings of the morphemes?
- (2) Do the morphemes have those meanings in other English words? (Don't use knowledge from foreign languages to "make the meaning work".)

22

□ "Abso-Bleedin'-Lutely"

(Quincy Jones, song in the film *Walk, Don't Run*, 1966)

□ "Well, I can *guaran-damn-tee ya*. Dannie's not playin'."

(Rick Reilly, *Shanks for Nothing*. Doubleday, 2006)

□ *Hallebloodyluah!*

□ In the movie *Wish You Were Here*, the main character expresses her aggravation (at another character's trying to contact her) by screaming *Tell him I've gone to Singabloodypore!*"

□ (George Yule, *The Study of Language*, 3rd ed. Cambridge University Press, 2006)

## Classification of Morphemes

20

(1) **Free Morpheme:** Can occur on its own. Ex: cat, tree, love, kick, the, was, yesterday etc.

Types of Free Morphemes

□ **Content Word** (open class, you can make up new ones): express general referential or informational content, a meaning that is essentially independent of the grammatical system of a particular language: *play, love, tree, geek, glasnost, novella* etc.

□ **Function Word** (closed class, you can't make up new ones): A free morpheme expressing syntactic relationships between units in a sentence, or obligatorily marked categories such as number or tense. : *it, was, he, my, the, and, under, who, a etc.*

(2) **Bound Morpheme:** Cannot occur on its own. Instead, it must be attached to something else. Ex: *un-, anti-, -ment, -s, -ed, etc.*

Compare:

I laughed yesterday

Yesterday I laughed

\*ed I laugh yesterday

23

□ *awfully: beawfullyware*

□ *bally: absoballylutely*

□ *bleeding: absobleedinglutely*

□ *bleep: fivebleepmile*

□ *blessed: absoblessedlutely*

□ *bloody: railbloodyway, whatsobloodyever, kangabloodyroo*

□ *blooming: absobloominglutely*

□ *damn: abdamnsurd, Piccadamnlilly*

□  *fucking: somefuckingplace, selfdefuckingfence*

□  *goddamn: ingoddamconsistent, guarangoddamnee*

□ *jolly: theojollylogical*

□ *posi: absoposilutely*

## Types of Bound Morphemes

21

□ **Non-Affixal (restricted):** Similar to a content word-- has a concrete meaning. However, cannot stand on its own. Ex: -scribe, as in *inscribe, prescribe*, etc. or -mit as in *permit, remit, submit*, etc., -kempt, -shevelled etc.

□ **Affixal:** A bound morpheme that has either grammatical meaning (similar to a function word), or modificalional meaning (including changing the lexical category of the thing you attach to).

## More English Reduplication

24

□ **Rhyming reduplication:** hokey-pokey, razzle-dazzle, super-duper, boogie-woogie, teenie-weenie, walkie-talkie, hoity-toity, wingding, ragtag, easy-peasy, tittle-tattle etc.

□ **Exact reduplications:** bye-bye, choo-choo, night-night, no-no, pee-pee, poo-poo.

□ **Ablaut reduplications:** bric-a-brac, chit-chat, criss-cross, ding-dong, jibber-jabber, kitty-cat, knick-knack, pitter-patter, splash-splash, zig-zag, flimflam. In the ablaut reduplications, the first vowel is almost always a high vowel and the reduplicated ablaut variant of the vowel is a low vowel.

□ **Shm-reduplication** can be used with most any word: e.g. *baby-shmaby, cancer-schmancer* and *fancy-schmancy*. This process is a feature of American English from Yiddish (began with the American Jews of New York City, then the New York dialect and then the whole country).

### 25 Where can/do affixes attach?

- prefix – attaches to beginning of word (un-like)
- suffix – attaches to the end of word (danc-ing)
- infix – inserted into the middle of a word (abso-bloody-lutely)
- Bontoc (Philippines)  
fikas 'strong'  
fumikas 'he is becoming strong'
- circumfix – part of the morpheme on each side of a word (en-light-en, a-woo-ing)
  - French negative: ne-suis-pas
  - Chikasav: chokma 'good' > **ɪkchokmʊ** 'not good'
  - German: getrunken 'drunk' + past participle
  - English: Gather ye rosebuds while ye may, Old time is still a-fly-ing
- Particular kind of affix, reduplicative (can be a prefix, or an infix), partial or full reduplication
  - Tohono O'odham plural affix= CV goks "dog" > gogoks "dogs"
  - Vedic Sanskrit, perfect verb (ja-gam, bi-bhed, ca-kar, ba-bhu, ta-tap)

## Tests to determine affix type

### 26 Does it change the part of speech?

- Derivational morphemes usually change the part of speech:
 

slow	adjective	slow-ly	adverb
write	verb	writ-er	noun
- Inflectional morphemes never change the part of speech:
 

dog	dog-s	both nouns
jump	jump-ed	both verbs

### 27 Is the focus within a word or between words?

- Derivational morphemes give semantic content to a word but do not indicate relationships between words.
 

un-like	flammable	pre-history
---------	-----------	-------------
- Inflectional morphemes express syntactic or semantic relations between words. They are generally required by the syntax.
 

two dog-s	John	jump-s
-----------	------	--------

## A Diversity of Morphemes

28

- Russian: -ina suffix means "meat"  
olen "reindeer" olenina "reindeer meat"
- Czech: -vaa- infix means "now & then"  
pratsvat "work" pratsvovavat "work now & then"
- Maasai: -u suffix means "inceptive" action  
aibor "I am white" aiboru "I become white"
- Sierra Miwok: the -i suffix means "distributive"  
po'al "to slit open" po'ali "to slit open several"
- Amharic: bale- prefix means "owner of"  
sue "store" balesuq "owner of a store"
- Nganasan: -zhay suffix means "toy variant"  
dinte "bow" dintezhay "toy bow"
- Polish: -ak means "young animal" (-ette)  
kot "cat" kochak "kitten"

## Tests to determine affix type (contd.)

29

### 3. How productive is the process?

- Derivational morphemes are usually limited in what they apply to (i.e., less productive).
 

flame + -able = flammable
tooth + -able = toothable??
un-interesting = uninteresting
un-slow = * unslow
- Inflectional morphemes apply to every word in a given lexical class (part of speech).
 

book-s	cat-s	wuk-s	gloof-s
--------	-------	-------	---------

### 4. How is it ordered?

- Derivational morphemes occur closer to the root than inflectional morphemes.
 

writ-er-s	govern-ment-s
-----------	---------------

### Practice on English

- Take 2 minutes to break the following words into morphemes. For each morpheme, decide if it is free or bound. For each bound morpheme, decided if it is derivational or inflectional.

baker's rejoined sillier thickeners unspeakably

## Types of Affixes: Inflectional Affixes, Derivational Affixes

27

### Derivational Morphemes

- Morphemes that change the meaning or the part of speech of a word.  
un- in unaffected, re- in reactivate, pro- in pro-choice
- ment in judgment, -ly in quickly, -ful in hateful
- Lie close to the root word
- Do not apply freely, constrained in various ways

### Inflectional Morphemes

- Serve a purely grammatical function, no change in primary meaning or category
- Never create a new word but create a different form of the same word - bushes, eaten, faster, John's
- Lie further from the root – close out derivations
- Apply across the board, less constrained

## a- and an- of Hindi

30

अनारुद्ध

अन्यंत

शांति shanti (peace)

शुभ shubh (auspicious, good)

शिक्षित shixit (educated)

शुद्ध shuddh (pure)

आरुद्ध

अत्यंत

अशांति

अशुभ

अशिक्षित

अशुद्ध

## Productivity, allomorphy etc.

31

- In word formation, as in phonology, we need to distinguish those things which we need to memorize about words on a word-by-word basis, and those things which we do not need to memorize on a word-by-word basis, because they follow from simple word formation rules.
- If the *word forms* that a morpheme can create are quite regular and allows the creation of new word forms, then the morpheme is said to be ***productive***.
- If the *word meanings* that a morpheme can create are quite regular, and do not need to be memorized on a word-by-word basis, then the morpheme is said to be ***predictable***.

## Order of Affixation, Ambiguities

34

Unhappiness, Noun

Prefix *un-*  
Stem *happy*  
Suffix *-ness*

**[un-[[happy<sub>ADJ</sub>] +ness<sub>NOUN</sub>] NOUN]**

OR

**[[un-[happy<sub>ADJ</sub>] ADJ]ness<sub>NOUN</sub>]**

What does un- prefix, nouns or adjectives?

## Productive Affix

32

- An affix that not only exists in the language, but which also has an active rule in the language that allows it to combine with other morphemes is a productive affix.  
The English -ed affix is productive: VERB + -ed = VERBpast  
The English -ing suffix is productive: VERB + -ing = Vparticiple
- An affix that exists in a language, but lacks an active rule that allows it to combine with morphemes to make new words is a non-productive affix.  

king-dom	play-wright
fief-dom	wheel-wright
wis-dom	cart-wright
serf-dom	ship-wright
martyr-dom	cart-wright

## Multiple Rules Generate Ambiguity, Different Representations

35

Order 1:

Step 1: VERB + -able = ADJECTIVE "capable of being VERBed"  
lock + -able = lockable "capable of being locked"  
Step 2: un- + ADJ = ADJ "opposite of ADJ"  
un + lockable = unlockable "opposite of lockable"

Order 2:

Step 1: un- + VERB = VERB "reverse the action of VERB"  
un + lock = lock "reverse the action of locking"  
Step 2: VERB + -able = ADJ "capable of being VERBed"  
unlock + -able = unlockable "capable of being unlocked"



## Types of Limitations

33

We already know what morphemic rules have to specify. They can also mention additional information, like combinational restrictions:

1. lexical category restrictions: re- can only attach to verbs
2. -ate usually only combines with Latinate stems
3. sound/ size restrictions:

Restrictions on the English comparative/superlative affix -er and -est

Restrictions on -en suffixation

## Rule-governed creativity

36

Creation of New Words, including:

- new combinations of morphemes (demosquitofy)
- invented words (xerox, teflon, nylon)
- Circumlocutions (cf. "Recursion") "a friend of a friend", "to get rid of all mosquitos"

New Words: Morphemic Combinations Some (rule-governed) possible morpheme combinations don't exist (cf. phoneme sequences like *blick*).

1. No one has made that combination before (demosquitofy)
2. There's already a word for it (\**to hospital* is not possible because *hospitalize* exists)
3. The rule for that morpheme is no longer active \**pot-wright*)
4. The rule for that morpheme sets specific limitations on what it can combine with (category, phonology, semantic etc.)

## Rules for Combining Affixes

37

Morphological Rules must state three things:

1. the lexical category an affix attaches to
2. where the affix attaches (prefix, suffix, etc.)
3. the lexical category of the resulting word  
 NOUN + -s = NOUN 'plural of NOUN'  
 ADJECTIVE + -en = VERB 'cause to be ADJ'  
 VERB + -s = VERB '3rd sing. of VERB'

Note: the different -s affixes of English look the same, but have different rules.

Present -s attaches to V → V <3 person present> *walk-s*

Possessive -s attaches to N → possessor *John's*

Plural -s attaches to N → N <many> *cat-s*

## Meaning Variation in a Single Morpheme

40

intermit (put between)	transmit (send across)
admit (let in)	remit (send back)
submit (hand in)	emit (to put forth)
permit (allow)	omit (to leave out)

\*vomit

impermissive/-ion, submissive/-ion, admissive/-ion,  
 remissive/-ion, transmissive/-ion, emissive/-ion,

permissive/-ion, omisive/-ion etc. but not vomissive/-ion

□ -mit seems to have a meaning involving change in location

for object or data (put/ hand/ send. etc.)

□ French equivalent is the verb *mettre*), but the output depends on the prefix. (Probably due to the fact that these words are all borrowed from Latin, at different times.)

## Cranberry Morphemes

38

A morpheme with clear meaning, but which only occurs in one word.

- blueberry - type of berry that's blue (blue + berry)
- cranberry - type of berry that's cran
- raspberry - type of berry that's rasp

Obviously, "berry" means something (i. e., BERRY), and "cran-" and "rasp-" seem to mean something, denoting a particular berry

- By extension, "cran-" also means "cranberry"

Ex: cranapple = cranberry plus apple

This only happens when the morpheme in question (i. e., cran-) is unique to some word.

(This is not the same as *rehearse*, where neither re- nor -hearse means rehearse all on its own.)

## Predictability and Productivity revisited

41

A + ize 'make (something) A'

- commercialize
- modernize
- standardize
- relativize
- prioritize
- criminalize

But...

- \*pleasantize
- \*sickize
- \*worthyize
- \*terribilize

## Portmanteau Morphemes

39

A morpheme that carries more than one piece of meaning, but which cannot be broken down into separate morphemes:

- English -s means "singular", "present", "3rd person", "indicative"
- *is* and *was*, the verb 'be', tense, person, number etc.
- Russian -u means both "feminine" and "accusative" at the same time.
- Hindi: hū <be+first person+sg+present>

These meanings cannot be separated into separate morphemes.

(Compare: Russ. -li means "past tense" and "plural", but it can be broken down: in this language -l- consistently means "past" and -i consistently means "plural".)

## In- prefix

42

- impossible
- immovable
- infallible
- invaluable
- intolerable
- indefinite
- irresponsible, irreverant
- illegal
- insane
- inoperable, inadvisable
- incongruous
- \*in-inflammable, in-intelligent, in-inflected

## Partially productive, and only partially predictable

43

V + N "N that Vs"

- tug-boat: boat that tugs
- cry-baby: baby that cries
- strong-box: box that is strong
- watch-dog: dog that watches
- light-house: house that lights
- hit-man: man that hits
- kill-joy: \*joy that kills
- break-water: \*water that breaks
- task-force: \*force that tasks
- spend-thrift: \*thrift that spends
- \*teach-book: book that teaches
- \*suck-course: course that sucks

## Some unpredictability

46

In many cases, the same kind of derivational pattern shows differences in form; take e.g. verb --> noun:

1) -al	refuse	refus-al
2) -ion	arrive	arriv-al
	confuse	confus-ion
3) -ation	extend	extens-ion
	derive	derivation
4) -ment	confirm	confirm-ation
	confine	confine-ment
	treat	treat-ment

This is in a sense *allomorphy*: the form of the nominalizing affix is something that depends on what host the affix is attached to (put differently, the different affixes only attach to certain hosts)

## Diseases, especially inflammatory ones -itis

44

▪ appendicitis	hepatitis
▪ tonsillitis	arthritis
▪ bronchitis	meningitis
▪ bursitis	dermatitis

Based on these forms and the *novel* forms below, we can conclude that the suffix '-itis' in English appears to be productive, although only in the narrow domain of disease-names.

- Monday-morningitis
- test-ititis
- olive-ititis

## Competition between morphemes

47

### Generalization:

- (1) -ity and -th are in competition with one another, complementary distribution  
 (2) -ness is generally available, and does not compete with -ity or -th

	-ity	-th	-ness
sincere	sincerity	sincerth	?sincereness
chaste	chastity	*chasteth	?chasteness
scarce	scarcity	*scarceth	scarceness
curious	curiosity	*curiousth	curiousness
warm	*warmity	warmth	warmness
deep	*deepity	depth	deepness
wide	*widity	width	wideness
red	*reddity	*redth	redness
sick	*sickity	*sickth	sickness
pleasant	*pleasantity	*pleasanth	pleasantness
happy	*happity	*happith	happiness

## Describing People

45

anarchist	pianist
chartist	violinist
communist	saxophonist
leftist	guitarist
Marxist	
feminist	linguist
Calvinist	racist
sado-masochist	sexist
*drummist	age-ist
*Lutherist	size-ist
	separatist
	unificationist

- The suffix '-ist' seems to be productive in English, based on the fact that we can easily coin novel words using the '-ist' suffix. However, the suffix '-ist' has multiple meanings corresponding to
  - (i) "advocate of"
  - (ii) "user of/expert in"
  - (iii) "detractor of"
- Although use (iii) is probably only a recently developed use of this morpheme, it appears to be the most productive.

## How is mental dictionary organized?

48

**Standard** dictionaries: alphabetical order, following spelling (not sounds)

- sieve, sift, sigh, **sign**  
 "sign" wouldn't be anywhere near other words that have the same sounds at the beginning, e.g. "psychology", "cycle"

**Rhyming** dictionaries: organized according to sounds at end of words

- **sign**, line, decline, resign, confine

**Thesaurus**: organized according to meaning

- **sign**: symbol, emblem, badge, mark, flag

## How is mental dictionary organized?

49

### Tip-of-the-tongue states

- "a navigational instrument used in measuring angular distances, especially the altitude of sun, moon and stars at sea" (Brown & McNeill 1966)
- subjects were asked to ...
  - guess the initial letter?
  - what other words come to mind?
  - what words rhyme with it?
- if they were unable to retrieve the word directly, some things they might nevertheless know are:
  - starts with /s/
  - sounds like "secant", "sextet"
  - meaning related to "compass"

## Open vs. Closed class: neurological evidence

52

### Broca's aphasia

- Production:  
"Well ... front ... soldiers ... campaign ... soldiers ... to shoot ... well ... head ... wound ... and hospital ... and so ..." (soldier describing how he was wounded)
  - Comprehension:
    1. He showed her baby pictures. (ambiguous for all English speakers)
    2. He showed her baby the pictures.
    3. He showed her the baby pictures.

(2 and 3 are unambiguous for normal English speakers, ambiguous for many Broca's aphasia patients)
- (Heilman & Scholes 1976)

## Anomia

50

"inability to retrieve words from mental dictionary"

### Category-specific anomias

- a. animate vs. inanimate
  - inanimate: ok (wheelbarrow, submarine, umbrella, briefcase, compass)
  - animate: impaired (wasp, duck, holly, ostrich, snail)
- b. fruits and vegetables, or furniture terms selectively impaired...

## Mental dictionary

53

### Contains

- morphemes
- rules for combining morphemes
- non-productive and non-predictable complex words

### Organization:

- not organized in single list
- organized according to phonological and semantic properties

## Priming

51

Task: Is this a word in English? Press button for "yes" or "no" as quickly as possible.

- NURSE
- DRAGE
- CAT
- TEF
- DOCTOR ... NURSE
- DOG .... CAT

Looking up a word in the mental dictionary **primes** semantically related words. i.e. People are faster to judge that NURSE is a word of English if they have just read DOCTOR. This is understood to mean that looking up DOCTOR in the mental dictionary makes it easier to look up NURSE. Similar finding for DOG and CAT and thousands of other related words.

Priming is also seen for phonologically related words

## Summary of Morphology

54

### Mental Lexicon and its organization

#### Morphemes

##### Kinds of morphemes

- roots and affixes
- bound vs free
- affix classification based on location (pre, suff, circum, in etc.)
- affix classification based on type (inflectional and derivational)

##### Word formation rules (WFRs)

- Productivity
- Predictability

##### How to isolate morphemes in languages

- sound-meaning pairing
- Influence of phonology on morphology
  - morpho-phonological rules
  - constraints on WFRs (word length, final segment)

## Nominal Inflection - Gender

class	semantics	prefix	singular	translation	plural	translation
1, 2	persons	m-/mu-, wa-	mtu	person	watu	persons
3, 4	trees, natural forces	m-/mu-, mi-	mti	tree	miti	trees
5, 6	groups, aug.	Ø/-j-, ma-	jicho	eye	macho	eyes
7, 8	artefacts, dim.	ki-, vi-	kisu	knife	visu	knives
9, 10	animals, loanwords, other	Ø/m-, Ø/n-	ndoto	dream	ndoto	dreams
11, 12	extension	u-, Ø/n-	ua	fence, yard	nyua	fences
14	abstraction	u-	utoto	childhood	-	-
singular		plural				
<b>mtoto</b> mmoja anasoma child one is reading One child is reading		<b>wtototo</b> wawilli wanansoma children two are reading Two children are reading				
<b>kitabu</b> kmoja kinatoshia book one suffices One book is enough		<b>vitabu</b> vitwill vitnatosha books two suffice Two books are enough				
<b>ndizi</b> moja zinatoshia banana one suffices One banana is enough		<b>ndizi</b> mbili zinatoshia bananas two suffice Two bananas are enough				

## Finnish Verb Inflection - Tense

indicative mood								
active voice	present tense		imperfect		perfect		pluperf.	
per. no. pron.	affirmative	negative	affirmative	negative	affirmative	negative	affirmative	negative
1st	minä puhun en puhu		puhun en puhunut		olen puhunut en ole puhunut		olin puhunut en ollut puhunut	
2nd sg. sinä	puhut et puhu		puhut et puhunut		olet puhunut et ole puhunut		olit puhunut et ollut puhunut	
3rd	hän puhuu ei puhu		puhuu ei puhunut		on puhunut ei ole puhunut		oli puhunut ei ollut puhunut	
1st pl.	me puhumme emme puhumme		puhunneet emme puhunneet		olemme emme ole olimme emme olleet puhunneet		puhunneet ette olleet puhunneet	
2nd pl.	te puhutta ette puhu		puhutta ette puhunneet		olitte puhunneet ette ole puhunneet		olitte puhunneet ette olleet puhunneet	
3rd	he puhuvat eivät puhu		puhuvat eivät puhunneet		ovat puhunneet eivät ole puhunneet		olvat puhunneet eivät oleet puhunneet	
passive voice		puhutaan ei puhuta puhuttai ei puhuttu		on puhuttu ei ole puhuttu		oli puhuttu —		

## Nominal Inflection - Case

	Singular	Dual	Plural
<b>Nominative</b> (Karttä)	-s,-s (-m,-m)	-äsi -au (-äk -i)	-äss -as (-äk -i)
<b>Accusative</b> (Karman)	-äss -am (-m,-m)	-äsi -au (-äk -i)	-äss -as (-äk -i)
<b>Instrumental</b> (Karaga)	-ää -ä	-ä-ga-ä -bhyaäm	-ä-ga-s -bhisis
<b>Dative</b> (Samprädäna)	-ä -e	-ä-ga-ä -bhyaäm	-ä-gas -bhayas
<b>Ablative</b> (Apäädäna)	-äss -as	-ä-ga-ä -bhyaäm	-ä-gas -bhayas
<b>Genitive</b> (Sambandha)	-äss -as	-ä-ga-s -os	-ä-ga -ämm
<b>Locative</b> (Adhikaraga)	-ä -i	-ä-ga-s -os	-su -su
<b>Vocative</b>	-s,-s (- -)	-äsi -au (-äk -i)	-äss -as (-äk -i)

## Other moods

conditional mood									
active voice	present tense		perfect		present tense		perfect		
per. no. pron.	affirmative	negative	affirmative	negative	affirmative	negative	affirmative	negative	
1st	minä puhuisin en puhuli		olisin puhunut en olisi puhunut		puhunen en puhune		lienen puhunut en liene puhunut		
2nd sg.	sinä puhuisit et puhusi		olisit puhunut et olisi puhunut		puhunet et puhune		lienet puhunut et liene puhunut		
3rd	hän puhuisi ei puhusi		olisi puhunut ei olisi puhunut		puhunee ei puhune		lienee puhunut ei liene puhunut		
1st	me puuhisimme emme puuhisi		olismme emme olisi puuhiseet		puhunemme emme puuhiseet		lienenme emme liene puuhiseet		
2nd pl.	te puuhisitte ette puuhusi		olisset puuhiseet ette olisi puuhiseet		puhunette ette puhune		lienet puhunette ette liene puuhiseet		
3rd	he puuhisivat eivät puuhisi		olivat puuhiseet eivät olisi puuhiseet		puhunevat eivät puuhune		lienevät puhunevat eivät liene puuhiseet		
passive voice		puhuttain ei puhuttasi		olisi puhuttu ei olisi puhuttu		puhuttaneen ei puhuttane ei puhuttane		lienee puhuttu ei liene puhuttu	

Finnish cases				
Case	Suffix	English prep.	Example	Translation
Grammatical				
Locative (internal)				
inessive (nessiliv)	-ssa	in	talossa	in (a) house
elative (elativ)	-sta	from	talosta	from (a) house
ilative (illativ)	-an, -en, etc.	into	taloon	into (a) house
Locative (external)				
adessive (adessiliv)	-lla	at, on	talolla	at (a) house
ablativ (abativ)	-ta	from	talotta	from (a) house
allative (allativ)	-lle	to	talolle	to (a) house
Essive				
essive (essiliv)	-na	as	talona	as a house
(excessive, dialectal) (eksessiliv)	-nta	from being	talonta	from being a house
translative (translativ)	-ksi	to (role of)	taloksi	to a house
Marginal				
instructive (instruktiv)	-n	with (the aid of)	taloin	with the houses
abessive (abessiliv)	-lla	without	talolla	without (a) house
comitative (komitativ)	-ne	together (with)	taloneen	with the house(s)

## Sanskrit – verb forms

45. Class 2 (root class), dviṣ 'hate', parasmaipada/active endings

	Present	Imperfect	Imperative	Optative
Sing.	1 dveṣmi	adveṣam	dveṣāṇi	dviṣyām
	2 dvekṣi	adveṭ	dviḍḍhi	dviṣyāḥ
	3 dveṣṭi	adveṭ	dveṣu	dviṣyāt
Dual	1 dviṣvah	adviṣva	dveṣāva	dviṣyāva
	2 dviṣṭah	adviṣṭam	dviṣṭam	dviṣyātam
	3 dviṣṭah	adviṣṭam	dviṣṭam	dviṣyātām
Plur.	1 dviṣmaḥ	adviṣma	dveṣāma	dviṣyāma
	2 dviṣṭha	adviṣṭa	dviṣṭa	dviṣyāta
	3 dviṣṭanti	adviṣṭan	dviṣṭantu	dviṣyūḥ

61

**46. Class 2 (root class), dviṣ 'hate', ātmanepada/middle endings**

	<i>Present</i>	<i>Imperfect</i>	<i>Imperative</i>	<i>Optative</i>
<i>Sing.</i>	1 dviṣe	adviṣi	dvesai	dviṣiya
	2 dvikṣe	adviṣṭhāḥ	dvikṣva	dviṣṭhāḥ
	3 dviṣte	adviṣṭa	dviṣṭām	dviṣṭita
<i>Dual</i>	1 dviṣvalhe	adviṣvahī	dvesāvahai	dviṣīvahī
	2 dviṣāt̄he	adviṣāt̄hām	dviṣāt̄hām	dviṣīyāt̄hām
	3 dviṣāt̄e	adviṣāt̄ām	dviṣāt̄ām	dviṣīyāt̄ām
<i>Plur.</i>	1 dviṣmahe	adviṣmahi	dvesāmahai	dviṣīmahi
	2 dviḍḍhve	adviḍḍhvam	dviḍḍhvam	dviṣīdhvam
	3 dviṣate	adviṣata	dviṣatām	dviṣīran

**Isthmus Zapotec**

1. palu	'stick'	7. spalube	'his stick'	13. spalulu	'your stick'
2. kubla	'dough'	8. skubabe	'his dough'	14. skubalu	'your dough'
3. tapa	'four'	9. stapabe	'his four'	15. stapalu	'your four'
4. geta	'tortilla'	10. sketabe	'his tortilla'	16. sketalu	'your tortilla'
5. bere	'chicken'	11. sperebe	'his chicken'	17. sperelu	'your chicken'
6. doño	'rope'	12. sto?obe	'his rope'	18. sto?olu	'your rope'

- a. Identify the morphemes that correspond to the following English translations:  
 possession (genitive)  
 3<sup>rd</sup> person singular  
 2<sup>nd</sup> person plural

- b. List the alternating pronunciations of the following:  
 'tortilla'  
 'chicken'  
 'rope'

What conditions the alternation in pronunciation?

62

**.2 Bontoc**

Bontoc is a language spoken in the Philippine Islands. Examine the data from Bontoc below and answer the questions that follow.

1. [fikas]	'strong'	5. [fumikas]	'he is becoming strong'
2. [kilad]	'red'	6. [kumilad]	'he is becoming red'
3. [batol]	'stone'	7. [bumato]	'he is becoming stone'
4. [fusul]	'enemy'	8. [fumusul]	'he is becoming an enemy'

- a. What word formation process is used to form the verbs? What part of speech class(es) can this process be applied to (items 1-4)?
- b. What type of affix is used to form the verbs? Describe its form and its placement within the root word.
- c. Given [pusi] 'poor', what would be the most likely meaning of [pumusul]?
- d. Given [njitad] 'dark', what would be the most likely form meaning 'he is becoming dark'?
- e. Given [pumukaw] 'he is becoming white', what is the most likely form meaning 'white'?

63

## From Morphology to Syntax

Word formation to phrase-structure

### More about Cranberry

Morphemes that have no meaning outside of the word they occur in

*It had been a rough day, so when I walked into the party I was very chalant, despite my efforts to appear gruntled and consolate. I was furling my wieldy umbrella... when I saw her...She was a very descript person...Her hair was kempt, her clothing shevelled and she moved in a gainly way.*

- *How I Met My Wife* by Jack Winter, *The New Yorker*, July 25, 1994

In the following groups of words, two words have the same suffix, one word has a different suffix, and one word has no suffix. Which ones go in which category?

- English has a prefix *dis*—:  
respect disrespect like dislike agree disagree please displease  
a. What is the word formation rule for *-dis*?
- b. How does the prefix *dis*— change the meaning of the word to which it is attached?

- Blueberry, blackberry, strawberry, boysenberry, loganberry, cranberry, mulberry, gooseberry
- Lukewarm, twilight, cobweb, hinterland
- Nonplussed, inept, ungainly, discern
- upholster (\*downholster), downhearted (\*uphearted), outlandish (\*inlandish)

Chamorro, a Malayo-Polynesian language spoken in Guam

[adda] 'mimic'	[aadda] 'mimicker'
[kanno] 'eat'	[kakanno] 'eater'
[tuge] 'write'	[tutuge] 'writer'
[atan] 'look at'	[atanon] 'nice to look at'
[saajan] 'tell'	[saajanon] 'tellable'
[guaija] 'love'	[guaiajon] 'lovable'
[tulaika] 'exchange'	[tulaikajon] 'exchangeable'
[chalek] 'laugh'	[chalekon] 'laughable'
[ngangas] 'chew'	[ngangason] 'chewable'
[nalang] 'hungry'	[nalalang] 'very hungry'
[dankolo] 'big'	[dankololo] 'very big'
[metgogot] 'strong'	[metgogot] 'very strong'
[bunita] 'pretty'	[bunitata] 'very pretty'

### Making new words

- Compounding  
bookworm, blackboard, blueberry, paper clip, sleepwalk
- Derivation: with affixes, zero derivation (father, knife, butter, queen, nail, ship, brush, referee), with stress changes (import, export, permit, conduct, reject, survey)
- Inflection
- Morpheme-Internal Changes (Ablaut)  
man~ men goose~geese mouse~mice louse~lice  
long~length wide~width strong~strength broad~breadth  
sing~sang~sung break~broke~broken strife~strive~striven,  
teeth~teethe slide~slid bid~bade ride~rode~ridden  
fly~flew~flown

## Semantic Shift: A shift from one meaning to an associated one.

*skirt an issue, he's green.* Often semantic shift is based on metaphor: conceptualizing one thing in terms of another - TIME IS MONEY, spend time, waste time, save time, etc.

*broadcast* originally meant "to cast seeds out"; with the advent of radio and television, the word was extended to indicate the transmission of audio/video signals. Outside of agricultural circles, limited use the word in the earlier sense *awful* originally meant "inspiring wonder (or fear)". It is a portmanteau of the words "awe" and "full", used originally as a shortening for "full of awe", awe-inspiring. In contemporary usage the word has an entirely negative meaning.

*gay* originally meant feelings of being "carefree", "happy", or "bright and showy"; it had also acquired some connotations of "immorality" as early as 1637. Later began to be used in reference to homosexuality, in particular, from the early 20th century, a usage that may have dated prior to the 19th century. In the early 21st century its connotation in American youth has shifted to reference something "stupid" or "silly".

A poet could not but be gay,  
In such a jocund company:

## Other sources of words

### Misanalysis and Backformation

Pea(se), cherry(cerise), orate (orator, oration), peddle (pedlar), burgle (burglar), edit (editor), enthuse (enthusiasm), liaise (liaison), resurrect (resurrection), preempt (preemption), televise (television), hamburger → chicken burger, veggie burger, sandwich → fishwich, bagelwich, Watergate → Irangate, Monicagate, Enrongate

### Eponyms

Fahrenheit, Celsius, Kelvin, sandwich, Argyle –socks, bougainvillea, boycott, boysenberry, caesarean, jumbo, gargantuan, paparazzo/i.

## Semantic change

- Narrowing:** Change from super-ordinate level to subordinate level. *skyline* any horizon to a horizon decorated by skyscrapers. *Hound* (dog), *meat* (food), *bird* (nestling, young bird), *litter* (bed, bedding, straw bedding for animals, waste), *accident*(any unforeseen event), *fowl* (any bird, wild fowl), *indigenous* (all inhabitants, native people), art (skill, aesthetic skill)
- Widening:** Change from subordinate level to super-ordinate level. *Kleenex*, *xerox*
- Metaphor:** Change based on similarity of thing, *mouse*
- Metonymy:** Change based on nearness in space or time, contiguity of concepts, e.g., *jaw* "cheek" → "jaw", the press, the crown, lend me your ear, palace, Washington.
- Synecdoche:** Change based on whole-part relation. Capital cities to represent countries or their governments, hands for labour, gray beard for old man, wheels for car, silver for flatware, threads for clothes etc.
- Litotes:** Change from weaker to stronger meaning, e.g., *kill* "torment" → "kill"
- Hyperbole:** Change from stronger to weaker meaning, e.g., *astound* "strike with thunder" → "surprise strongly".
- Degeneration:** e.g., *knav* "boy" → "servant" → "deceitful or despicable man".
- Elevation:** e.g., *knight* "boy" → "knight".

## Other sources of words

### Blends

smog, motel, brunch, chortle, frumious, infomercial, modem, cyborg

### Clipping

gym, gas, plane, fan, piano, van, ad, bike, math, phone, rad, fax, lab, dorm, prof, info

### Acronyms

NASA, UNICEF, snafu, radar, laser, scuba, ROM, IIT, PAF

### Coinage

Nylon, kodak, xerox, teflon, kleenex, jello, vaseline, nerd, quark

- Borrowing**  
gymkhana, shampoo, pajama Hindi  
spaghetti, mafia, ciao, Italian  
pistol, robot Czech,  
genre French,  
yak Tibetan  
kindergarten German,  
catamaran, rice, cheroot, ginger, mulgatawny Tamil  
sable mammoth glasnost perestroika Russian  
cot, brinjal, loot, cowrie, aniline, sulphur Sanskrit  
ambassador, embassy, vassal Gaulish (Celtic Language)

## Morphology and Syntax

- Morphological structure:** order of morphemes, word formation rules, computing meaning from the parts, hierarchical structure of morphemes in a word
- Some grammatical relations can be expressed either morphologically or syntactically
  1. Comparison of adjectives in English
  2. Possession
- Phrase structure:** Just as each language has rules for assembling morphemes into words, so each language has rules for assembling words into phrases and sentences. The rules of syntax are extremely productive: we do not learn a list of sentences, we learn how to construct and decode new ones.

## Syntax: knowledge of sentences

- Assembling words into phrases and sentences
- (a) What kind of rules are there in human languages for putting words together into phrases and sentences?
- (b) In any particular language, what are the rules for putting words together?
- Data
- (a) Patterns of grammaticality that cannot be reduced to meaning.
- (b) Aspects of meaning that do not come from the individual words (green grapes and apples, my aunt in France)
- (c) Similarities and differences between languages in these patterns.

## Alice's reaction

- 'It seems very pretty,' she said when she had finished it, 'but it's rather hard to understand!' (You see she didn't like to confess, even to herself, that she couldn't make it out at all.) 'Somehow it seems to fill my head with ideas--only I don't exactly know what they are! However, somebody killed something: that's clear, at any rate'.

## Jabberwock!



## Jabberwocky – Through the looking glass

'Twas brillig and the slithy toves  
Did gyre and gimble in the wabe;  
All mimsy were the borogoves,  
And the mome raths outgrabe

Beware the Jabberwock, my son!  
The jaws that bite, the claws that catch!  
Beware the Juju bird, and shun  
The furious Bandersnatch!"

He took his vorpal sword in hand:  
Long time the manxome foe he sought --  
So rested he by the Tumtum gree,  
And stood awhile in thought

And as in uffish thought he stood,  
The Jabberwock, with eyes of flame,  
Came whiffling through the tulgey wood,  
And burbled as it came!

(by Lewis Carroll, 1872)

## Judge the grammaticality of

- Sentences have Structure: One cannot simply put words in any order  
*The young man is reading the old book on the porch.*  
\**On the reading old is young porch man the the book.*
- Goals for Syntacticians
  - Describe the differences between grammatical and ungrammatical sentences.
  - Describe and explain the order of words and the structure of sentences.
  - Describe differences between related structures.
  - Explain ambiguity.
  - Determine the finite means by which we construct an infinite set of expressions.

## Grammaticality: What it's not

- Reminder: Prescriptive grammar and descriptive grammar are not the same things.
  - *Descriptive grammar* : grammar rules that the language centers of your brain use automatically.
  - *Prescriptive grammar* : grammar rules that you consciously learn as a form of "linguistic etiquette"
- We are concerned with descriptive grammar in this class, since it tells us what the brain does.

## Grammaticality Judgments Do Not Depend On...

- what you were taught in school: \* *It is I* is an **ill-formed** sentence of English, even though it might be taught in schools.
- familiarity with a sentence: *That's not my grandmother's racquet in your cat* is grammatical, even though you've never heard it before.
- Understandability: You can often understand an ill-formed sentence: *The book are in the cupboard* is understandable. It's just put together wrong.
- whether the sentence describes a factual situation: sentences like *Colorless green ideas sleep furiously* and *Tarzan is the Queen of England* are grammatical, but either meaningless or untrue.

- 
- The template could be modified to generate even more sentences by adding the idea of optional elements.
  - A element shown in parentheses is *optional* in this template:
  - (article)- noun-( aux. verb)- verb-( article)- noun
  - The modified template can now be used to generate even more sentences:  
*Foxes are eating grapes.*  
*The turtle will eat flowers.*  
*We could see the books.*

## Beads- On- A- String Model of Syntax

- Beads- on- a- string Model: A sentence consists of words that are placed in a specific linear order, like beads threaded onto a string. There are rules that tell you what order to use when putting the beads on the string.
- The sentence *The blue whale ate the plankton* might have the following structure:  
**THE →BLUE→WHALE→ATE→THE→PLANKTON**

- 
- You could even imagine that a language has several different templates that could be used for generating Ss. For example:
  - (article)- noun-( aux. verb)- verb-( article)- noun *or*
  - (article)- noun-( aux. verb)- verb
  - The first template is the one we've been using. The second one is new. The second one lets us make sentences like the following:
    - *The foxes are sleeping.*
    - *The books might fall.*
    - *Bats can fly*

- 
- What kinds of rules would be necessary to generate this structure? One type of rule that might work well would be a template. By following the template, you could generate different Ss:
  - Ex: **article- adjective- noun- verb- article- noun**
  - This template could be used to generate the following sentences, and many others:  
*The blue turtle ate the flowers*  
*A sly fox stole the fruit.*  
*The happy girl signed her name.*

## Characteristics of the Beads Model

- 
- Encodes only linear order. The only relations between beads (words) are "before" and "after". Each bead is a separate entity. There are no entities other than the words (beads) and the sentence as a whole (beads on string).
  - Problem: Even if we made enough templates to generate every possible sentence of a language, this model would be insufficient. This is because the beads- on- a- string model does not encode certain information that people seem to use when making/ understanding/ modifying or otherwise using sentences.

## What do we know about sentences?

- One things humans know about sentences is that they contain *subjects* and *predicates*. The beads model does not indicate what the subject is, and what the predicate is. Is this a problem?
- Consider subject- verb agreement. If the subject of a sentence is singular, the verb must also be singular. If the subject is plural, the verb must also be plural:  
The boy/he is in my class. (sing. *boy* = sing. *is*)  
The boys/they are in my class. (pl. *boys* = pl. *are*)

## Evidence for Constituents

- Usable in isolation:  
*What did it eat?* The flowers.  
*What did it do ?* Eat the flowers.  
*Who ate the flowers?* The blue turtle.  
*The what ate the flowers?* Blue turtle.  
*You can never use "ate the" in isolation...*

- Can we formalize the rule of subject- verb agreement using the template model? The only way to do this is to have separate templates for singular and plural sentences. This is because the idea of subject **cannot be defined in terms of linear order**.
- Ex: subject cannot be defined as the first noun in a sentence, nor as the noun closest to the verb:  
The *children* are scary.  
The *children* who are in Halloween costume are scary.  
    1 st sentence: subject = your children  
    2 nd sentence: subject = that your children play with matches
- The subject can be a *phrase*. The beads model does not possess *phrases* within sentences.

- Can be topicalized (clefted):  
It was *the flowers* that the blue turtle ate.  
It was *the blue turtle* who ate the flowers.  
It was *eat the flowers* that the blue turtle did.  
*You can never cleft "ate the" for emphasis:*  
*\*It was ate the that the blue turtle did flowers.*

## Phrases within Sentences

- In fact, there are a lot of places where the idea of phrases within sentences is helpful in understanding how we use sentences. These phrases are referred to as **constituents**.
- Going back to our original sentence *The blue turtle ate the flowers*, we have intuitions that there are several different constituents (sub- phrases) in this sentence:  
The blue turtle | ate the flowers (subject vs. predicate)  
even smaller constituents: blue turtle and the flowers
- Note: "ate the", "turtle ate", "blue flowers", etc. are not constituents

- Can be replaced by a placeholder/substitution:  
*It ate the flowers.* it = the blue turtle  
The blue turtle ate *them*. them = the flowers  
The blue turtle *did*. did = ate the flowers.  
*There is no placeholder for that can be used for "ate the".*

□ Can be conjoined:

*The blue turtle and the green fish ate the flowers.*  
*The blue turtles and green fish ate the flowers.*  
*The blue turtle ate the flowers and the seaweed .*  
*You cannot conjoin "ate the" with anything:*  
*\*The blue turtle ate the and enjoyed the flowers.*

## Nouns (N)

- Can be inflected for number (singular, plural), case, gender and other properties
- Can be modified by adjectives, demonstratives (*this/that*), articles (*a/the*)  
*Those big bulldozers \*those big ran*
- Semantically, generally denotes a person, place, thing or abstract concept

□ Can be moved:

If a sequence of words can be moved as a group, they may form a phrase:  
*The flowers, the blue turtle ate (but not the leaves)*

□ **Corollary** there are, frequently, several different ways of combining the same words to come up with the same basic meaning. When you do this, the words that form constituents tend to stick together:  
*The blue turtle ate the flowers.*  
*The flowers were eaten by the blue turtle.*  
*Did the blue turtle eat the flowers?*  
*The flowers were eaten by the blue turtle?*  
the blue turtle and the flowers seem to stick together.  
 But other combinations, like "eat the" get broken up.

## Verbs (V)

- Can inflect for tense, person, negative, etc.
- In English, the verb usually occurs after the subject and before the object:  
*Chris ate the pie*
- Generally denotes events, states, actions

## Lexical Categories

- The "Parts of Speech" that different words belong to
- You can determine the lexical class of a word
  - a) by the morphology it takes
  - b) by its syntactic position with respect to other word types
  - c) by its function in use
  - d) by its meaning

## Adjectives (Adj/A)

- May occur before a noun (*huge eyes*), or as a predicate modifying a subject (*The man is deaf*)
- May inflect for comparative/superlative (*larger, largest*) and be modified by adverbs (*very large*)
- Generally modify a noun

## Adverbs (Adv)

- Generally don't carry inflectional morphology
- Positioning may vary:  
*Robin slowly opened the door*  
*Robin opened the door slowly*  
*Slowly, Robin opened the door*
- Generally modify a verb, adjective or other adverb

## Determiners (D), Quantifiers (Q), Conjunctions (Conj), Pronouns (Pr)

- Determiners: this term covers a range of classes:
  - Articles: definite *the*, indefinite *a*
  - Demonstratives: *this/that*
  - Possessives: *my, your, Carmen's*
- Quantifiers: *Many, all, some, few, a lot, lots, hundreds,*  
 Numerals: *one, fourth*
- Conjunctions: Join words, phrases or sentences    *and, or, but, although, if*
- Pronouns: Replace a noun phrase  
*I, you, he, she, it, we, they* (subject pronouns)  
*me, him, her, us, them* (object pronouns)

## Open vs. Closed Lexical Classes

- Open lexical classes: one can easily add new members to via the word-formation processes previously discussed; have large numbers of members (nouns, verbs, adjectives, adverbs)
- Closed lexical classes: have limited membership, resist incorporation of new members (prepositions, articles, demonstratives, auxiliaries quantifiers etc.)

## Auxiliary Verbs

- Often indicate tense, aspect or modality, must co-occur with another (main) verb
 

*I must go.*  
*I will go.*  
*I do eat peanut butter sandwiches.*
- English: *has, have, had, do, did, does, shall, will, should, would, may, might, must, can, could, ought; is, am, are, was, were, be, being, been* (but not copula *be* and possessive *have*).
- **Note:** Lexical classes should be determined by their behaviour with respect to morphology and syntax, rather than by their meaning.

## Prepositions / Postpositions (Adpositions) (P)

- Prepositions and postpositions denote location, direction, and a range of other notions.
- Prepositions precede a noun phrase  
*To the house; from the store, with Wendy*  
 Postpositions follow a noun phrase  
 Finnish: *joen yli* river over "over the river"  
 Hindi: *billi ko, raam ne, ghar se*

## Constituents

- Sentences are made up of subparts, called "constituents"  
*The young man is reading the old book on the porch.*
- Constituents (units, subparts)  
 The young man || is reading the old book on the porch  
 || is reading the old book || on the porch  
 || is reading || the old book||

## Syntax

The rules of syntax must generate **all** and **only** the grammatical sentences of a language.

- The new mother slept the baby
- The new mother slept soundly
  
- Juliet believes Romeo to be a gentleman
- Juliet believes to be a gentleman
- Juliet tries Romeo to be a gentleman
- Juliet tries to be a gentleman
- Juliet wants Romeo to be a gentleman
- Juliet wants to be a gentleman

### Judge the grammaticality of...

- Jack and Jill ran up a large hill  
Jack and Jill ran up a large bill
- Jack and Jill ran a large hill up
- Jack and Jill ran a large bill up
- Up a large hill ran Jack and Jill
- Up a large bill ran Jack and Jill
- Walter put the sweater on  
put on the sweater  
put it on  
put on it

### Exercise: Identify the phrases

- (a) The boy stood on the train.
- (b) Sita gave the ring to Hanuman.
- (c) This teacher knew that the student had skipped the class.

- Tony gave a book to her.
- Tony gave her a book.
- Tony donated a book to her.
- Tony donated her a book.
  
- Caesar made Cleopatra leave
- Caesar let Cleopatra leave
- Caesar saw Cleopatra leave
- Caesar wanted Cleopatra leave

### Creativity

- $S \rightarrow NP VP$   
A sentence consists of a noun phrase followed by a verb phrase, the idea of subject and predicate.
- Noun Phrase  
Caesar, the cat, seven young llamas from Peru, Cleopatra, a feeling, the strangest love story that you have ever heard,
- Verb Phrase  
wept, fed the dog, ran through the fields, said that he thought that the large Hadron collider was unlikely to find anything, walked as if on eggshells

- NP + VP combinations
  1. Caesar wept.
  2. The cat fed the dog
  3. Seven young llamas from Peru ran through the fields.
  4. Cleopatra left
  5. A feeling said that he thought that the large Hadron collider was unlikely to find anything
  6. The strangest love story that you have ever heard walked as if on eggshells

## NP

I saw *the cat*.  
 It is *the cat* that I saw.  
 I saw *it*.  
 \*It is *the* that I saw *cat*.

This phrase is called NP (Noun Phrase), because the head of the phrase is N.

## Phrase structure/Rewrite rules

- Pieces of NP & VP
 

NP → Det N	NP → N
VP → V NP	VP → V
- How many Sentences with this tiny Grammar (five rules) and 3 determiners: the, four, some  
 3 nouns: dogs, cats, slugs  
 3 verbs: understood, ate, approached  
**468 sentences (12 NP x 3VP + 12NPx3VPx12NP)**
- 10 determiners, nouns, verbs: **122,100 sentences!**

## PP

I ate a sandwich *In this house*.  
*In this house*, I ate a sandwich.  
 I ate a sandwich *there*.  
 \**In*, I ate a sandwich *this house*.  
 \**In this*, I ate a sandwich *house*.

This phrase is called PP (Prepositional Phrase), because the head of the phrase is P.

## Phrase Names

- Thus, a sentence consists of multiple phrases, and one phrase is nested with another. The names of the phrases use the idea of categories, and the distributional definition of categories.
- **Phrases are named after their HEAD:**
  - The first element of NP is N. (e.g. men with a long hair)
  - The first element of PP is P. (e.g. in the park)
  - The first element of VP is V. (e.g. ate the cake)
  - The first element of AP is A. (e.g. very tricky)

**Principle: Every word forms a phrase on its own.**  
 Recall the constituency tests that you can apply.

## VP

The man ate a sandwich.  
 The man ate a sandwich, and Fred did too. (did too = ate a sandwich)  
 It was eat a sandwich that Fred did.  
 This phrase is called VP (Verb Phrase), because the head of the phrase is V.

## AP

The President was *quite firm* that she did not want BT Brinjal.

*Quite firm* was the President that she did not want BT brinjal.

It was *quite firm* that the President was that she did not want BT brinjal.

This phrase is called AP (Adjective Phrase), because the head of the phrase is Adj.

## Constituents broken up - Heads

All phrase types are built around a member of the same category; i.e., Nouns build NPs, Verbs build VPs, Adjectives build AP etc. The category around which a phrase is constructed is called the **HEAD**, as we have just seen.

Phrases may occur with only a filled HEAD, so John is a Noun and an NP, and walk is a verb and also a VP etc.

The HEAD word may occur with other elements

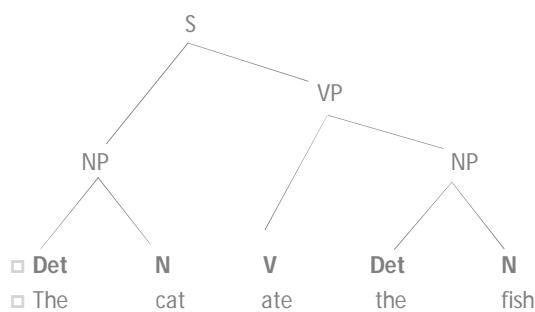
- (a) Specifier
- (b) Complement

## From rules to trees

- An alternative way to represent constituent structure is through tree diagrams. There are two ways to draw: bottom-up and top-down. The procedure of drawing a tree (Bottom-up)
  1. Write words in a sentence (or words in a part of the sentence).
  2. Write the categories of the words (e.g. N, V, P, Det) above them.
  3. Write the phrase of the category (e.g. NP, VP, PP) above it. (Start from the rightmost word.)
  4. Group the members of a phrase together, using the rules above.
  5. Repeat (iii) and (iv). (From right to left)
  6. Write "S" at the top, and group NP and VP together, using lines.

## Specifiers

- Semantically, specifiers help make the meaning of the head more precise. Syntactically they mark the phrase boundary – in English specifiers occur at the left boundary of the phrase, daughters of the phrasal node
- Determiners are NP specifiers
- Degree words like *almost*, *quite* and *very* are Adj and P specifiers
- Subjects actually begin as Specifiers of VP (consider that the role played by the subject reflects a property of the verb; but we won't worry about this for now)



## Complements

- These provide information about entities and locations whose existence follows from the meaning of the head. The verb *beat* requires something to be beaten, and the preposition *in* requires a location and the noun *argument* requires someone to argue with and a matter for argument.
- Phrases now contain three kinds of objects; Heads, Specs and Complements
- Generalized Phrase Structure Rule for any head X in English:  $XP \rightarrow (\text{Spec } X \text{ (Comp) } )$

- Very fond of all cheeses
- Told the boy that the book lay on the table
- The children in their uniforms
- Down into the cellar of the restaurant
- Walk very quietly into the auditorium
- The students of syntax
- The children in their uniforms walked very quietly into the auditorium

## More verbal subcategorization

<input type="checkbox"/> Null	vanish, die, arrive	<i>The child arrived.</i>
<input type="checkbox"/> NP	devour, cut, prove	<i>The lion devoured the deer.</i>
<input type="checkbox"/> AP	be, become	<i>The man became angry.</i>
<input type="checkbox"/> PP(to)	talk, refer, dash	<i>The dog dashed to the door.</i>
<input type="checkbox"/> NP NP	spare, hand, give	<i>We handed the man the map.</i>

## Explaining grammaticality

- The phrase structure rules of English (we have developed so far) can generate ungrammatical sentences, e.g.
  - \*The man loves.
  - \*The boy built.
  - \*The bus arrived the station.
  - \*They saw me the house.
- To deal with this over-generation problem, we need to make the grammar pay attention to the requirements of the particular Head in each sentence. Some verbs require object NPs, some don't. Some verbs require PPs, some don't.
- The solution: In the lexicon, we add information about what each head requires or allows to occur with it, i.e., their complements.

<input type="checkbox"/> NP PP (for)	buy, cook, reserve
<i>We bought a hat for the priest.</i>	
<input type="checkbox"/> NP PP (to)	hand, give, send
<i>He gave a diploma to the student.</i>	
<input type="checkbox"/> NP PP (loc)	put, place, stand
<i>He stood the snowman near the gate.</i>	
<input type="checkbox"/> PP (to) PP (about)	talk, speak
<i>I talked to the students about the elections.</i>	
<input type="checkbox"/> NP PP (for) PP (with)	open, fix
<i>We fixed the lock for the priest with a screwdriver.</i>	

## Subcategorization

- put [\_\_ NP PP] (*put* requires an NP and a PP)  
They put the car in the garage.
- sleep [\_\_] (*sleep* does not take any NP)  
John slept. /\*John slept the car.
- Many verbs are classified into more than one frame:  
e.g. give [\_\_ NP NP] and give [\_\_ NP PP]  
eat [\_\_ NP] and eat [\_\_]

## Noun complements

<input type="checkbox"/> Null	car, boy, electricity
<input type="checkbox"/> PP(of)	memory, failure, death
<input type="checkbox"/> PP(of) PP (to)	gift, description, donation
<input type="checkbox"/> PP(with) PP (about)	argument, discussion, conversation

## Adjective complements

- Null tall, green, smart
- PP (about) curious, glad, angry
- PP (to) apparent, obvious
- PP (of) fond, full, tired

## Sentence as complements contd.

- Nouns claim, belief, fact, knowledge, proof, conclusion  
*They lack proof that the earth is flat.*
- Adjectives afraid, certain, aware, confident  
*They are confident that the earth is flat.*
- Prepositions over, about  
*They argued about whether the earth was round.*

## Preposition complements

- Null near, away, down
- NP in, on, by, near, down
- PP down, up, out

*He walked out of/down into the cellar.*

## Summary

- The lexical properties of a head restrict the kinds of phrase structures in which they can participate.
- The elements that occupy Spec and Comp positions are themselves PHRASES.
- This RECURSION is a fundamental property of all human languages and is what permits **infinite output from finite means**.

## When Complements are Sentences

- Verbs believe, know, think, remember (S)  
*They believed that the earth is flat.*
- persuade, tell, convince, promise (NP S)  
*They persuaded Galileo that the earth was flat*
- concede admit (PPto S)  
*They conceded to Galileo that the earth was round.*

## Exercise

1. How many subcategorization frames for each of the following verbs? Write out the frames and give real-sentence examples.  
send think deny want
2. Are there verbs that have the following subcategorization frame?  
[ \_\_ NP S]

## Complement clauses again

- The sentential complements are preceded by certain words as is shown below:

The psychic knows *if* the contestant will win

*that*

*whether*

These words are called Complementizers and they must be positioned in the tree.

CP → (Spec) C (Comp)

Specifier ?

Complement? The sentence [the contestant will win].

- The lexemes that select or permit clausal complements participate in a rewrite rule: XP→ Spec X (CP)

## Syntax contd.

### Complements vs. modifiers

- Obligatory phrases

- \*Tom fed.
- \*Anne hit.
- \*Minnie mentioned.
- \*Mickey put the book.
- \*Pluto devoured.
- Tom fed Jerry.
- Anne hit the wall.
- Minnie mentioned her favourite book
- Mickey put the book on the table.
- Pluto devoured the bone.

- Optional phrases

- Cinderella sang a song (in the bathtub).
  - Anastasia slept (all morning).
  - The train arrived (at three o'clock).
- Time and location information PPs are usually optional to the sentence structure.

### When Complements are Sentences

- Verbs

believe, know, think, remember (S)

*They believed that the earth is flat.*

persuade, tell, convince, promise (NP S)

*They persuaded Galileo that the earth was flat*

concede admit (PPto S)

*They conceded to Galileo that the earth was round.*

### Arguments

- Dealing with the *obligatory requirements* of verbs (recall subcategorization information), the NPs, PPs, CPs etc. that a verb **must** combine with are known as the **arguments** of the verb. When one of the **arguments** of a verb is missing, the sentence becomes ungrammatical.

- Missing subjects

- \*fed the dog

- \*slept

- Missing objects

- \*Tom fed (Jerry)

- \*Tom gave Jerry (cheese)

- Nouns

claim, belief, fact, knowledge, proof, conclusion

*They lack proof that the earth is flat.*

- Adjectives

afraid, certain, aware, confident

*They are confident that the earth is flat.*

- Prepositions

over, about

*They argued about whether the earth was round.*

### Modifiers

- In addition to the arguments of a verb, there are other phrases that can be optionally included in a sentence and are called **modifiers**. The PP modifier below describes where the action described by the verb takes place.

The cat sat

The cat sat on the mat

- Modifiers occur recursively and with different kinds of VPs, another way of generating infinitely many sentences

The cat sat on the mat (intransitive V)

The cat sat on the mat in the morning

The cat sat on the mat in the morning during the summer

The cat chased the dog in the morning (transitive V)

The cat put the ball in the box in the morning (ditransitive V)

The cat realized [that the dog takes naps in the afternoon] in the morning  
(note that the PP "in the morning" modifies the VP realized and not take naps)

## Distinguishing between Arguments and Modifiers

### a. Obligatory phrases are arguments

Tom fed Jerry in the morning.  
Minnie put the book on the table after lunch.

### b. Implied by the 'core' meaning of verb

feed  
eat  
send  
think  
give

## NP modifiers vs. arguments

The two NPs below appear to have the same structure:

- This<sub>Det</sub> student<sub>N</sub> [of syntax <sub>PP</sub>]
  - This<sub>Det</sub> student<sub>N</sub> [in the elective class <sub>PP</sub>]
- This **student in the elective class** is better than that **one**  
 This **student** in the elective class is better than that **one** in the minor class  
 This **student of syntax** is better than that **one**  
 \*This **student** of syntax is better than that **one** of drama  
 The **one** pronominal replaces N and complements/arguments obligatorily and optionally any modifiers.

### c. do-so substitution

1. Tom **cooked lunch**, and Jerry did so (too).
  2. \*Tom **cooked** lunch, and Jerry did so dinner.
  3. Tom **put a book on the shelf**, and Jerry did so (too).
  4. \*Tom **put a book** on the shelf, and Jerry did so on the table.
  5. Tom **read a book** in the morning, and Jerry did so in the evening.
  6. \*Tom **read a book** in the morning, and Jerry did so a magazine in the evening.
  7. Tom **gave the cheese to Pluto** and Jerry did so too
  8. \*Tom **gave the cheese** to Pluto and Jerry did so to Mickey
  9. \* Tom **gave the cheese to Pluto** and Jerry did so the pie to Mickey.
  10. \*Tom **gave** the cheese to Pluto and Jerry did so the pie.
- The substitute **do-so** must replace the V and its arguments. However...

### • Another NP example

The claim that the earth is round is well established.  
 This claim that the earth is round is well established while that claim that the earth is flat is not.  
 This **claim that the earth is round** is well-established while that **one** is not.  
 \*This **claim that the earth is round** is well-established while that **one** that the earth is flat is not.

**Do-so** may replace modifers too along with the verb and its arguments

- a. The cat **chased the dog**, and the ferret did so too.
- b. The cat **chased the dog in the morning**, and the ferret did so too.
- c. The cat **chased the dog** in the morning, and the ferret did so in the afternoon
- d. The cat **chased the dog** in the park in the morning, and the ferret did so in the living-room in the afternoon
- e. The cat **chased the dog in the park in the morning**, and the ferret did so too.

## Other differences ...

### • Complements precede modifiers

The student of syntax in the elective class is from India

\*The student in the elective class of syntax is from India

### • Multiple modifiers possible

The student of syntax in the elective class, with red glasses...

### • Modifiers may be freely reordered with respect to each other but not complements

The student of syntax, with red glasses, in the elective class ...

### • Complements can be co-ordinated with complements and modifiers with modifiers but mixed co-ordination is ruled out

The student of syntax and of drama in the elective class.

The student of syntax in the elective class and with red glasses...

\*The student of syntax and with red glasses...

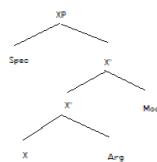
- **Extraposition:** It is possible to extrapose (move to the periphery) adjuncts PPs but not possible to extrapose complement PPs.  
A student of syntax/a student with red glasses came to talk to me yesterday.  
\*A student came to talk to me yesterday of syntax.  
A student came to talk to me yesterday with red glasses.
- **Preposing:** Complements can be preposed but not modifiers  
John is a student of syntax with red glasses  
[What area of linguistics] is John a student of?  
\*[What colour of glasses] is John a student with?
- Note that Complements and Adjuncts go in **opposite** directions with respect to Extraposition and Preposing.
- Heads are more closely related to their complements than to their adjuncts.

## Explaining Ambiguity

- An ambiguous sentence is a sentence that has two meanings. Some sentences are ambiguous because they contain a word that is ambiguous. Such cases are called **lexical ambiguities**.
- By contrast, some sentences are ambiguous without containing any ambiguous words. These cases can be explained when it is observed that the sentences in question can be given two distinct syntactic trees, leading to what is called a **structural ambiguity**.

## Phrase trees revisited

- The preceding slides suggest that there is a structural difference between arguments and modifiers. Arguments are direct sisters of heads and modifiers are sisters of head+arguments
- $XP \rightarrow (\text{Spec}) X$
- $X' \rightarrow X' (\text{Modifier}) \quad \text{additional stuff}$
- $X' \rightarrow X (\text{Complement}) \quad \text{required by the head}$



**John is sitting near the bank.**

- a. Meaning 1: bank = financial institution
  - b. Meaning 2: bank = slope at the side of a river
- This sentence has two meanings, which appear to be reducible to the two meanings of the word 'bank'.

The little boy is sitting next to a **bat**

This **ball** is extremely well-known in our town.  
Where is the **pitcher**?

## Sentence reconsidered

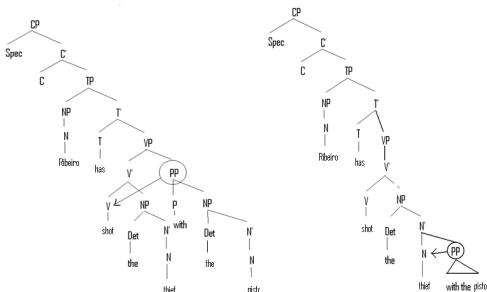
- The relation between subject and predicate is indicated through agreement (subject-verb or also called Spec-Head agreement)
- Information regarding the time and aspect of the event indicated by the verb are not confined to the verb but range over the whole sentence.
- Let's posit a node T that builds TPs which replace our Ss. Rewriting the PSR for S and CP we have:  
 $TP \rightarrow (\text{Spec}) T (\text{Comp})$ , Spec is NP and Comp is VP  
 $CP \rightarrow (\text{Spec}) C (\text{Comp})$ , where Comp is TP

Such a reduction is not possible in the following example, however:

**Ribeiro has shot the terrorist with the pistol**

- Meaning 1: The cop did the shooting with the pistol  
(PP modifier of V', shoot)
- Meaning 2: The cop shot (with something) the thief who had the pistol  
(PP modifier of N', thief)
- None of the words are ambiguous. So why is the sentence ambiguous?
- Because it can be given two distinct syntactic trees, given where the PP is attached.

## Attachment ambiguities



## Crosslinguistic variation

- **Lexical categories:**  
Nouns and verbs are found in all human languages but not others  
Example, in Korean adjectives are replaced by nouns or verbs  
**Ku chayk-i caymi issa**  
that book-nom interest exists  
*That book is interesting*  
**Ku pap-i mas issa**  
that food-nom taste exist  
**ku pap-i maypta**  
that food-nom be-spicy  
**cip-i khuta**  
house-nom be-big
- In Jacalteco, a Mayan language, prepositions are absent  
xto naj conob  
went he town  
*He went to town*  
ai naj s-wi' witz  
is he head-hill  
*He is on the hill*
- Nootkan languages have only N and V  
English does not have classifiers or preverbs found in other languages

## Much ambiguity – many parses

- Tom realized that Jerry eats the cheese in the bathtub on Thursday
- Possible readings include:
- (a) cheese is eaten in the bathtub and on Thursday (both modifiers are attached to the lower verb, eat)
  - (b) The realization that Jerry eats cheese came to Tom in the bathtub and on Thursday (both modifiers are attached to the higher verb, realize)
  - (c) Tom realized on Thursday that Jerry ate cheese in the bathtub (one modifier, closest, is attached to the lower verb eat and the other one to the higher verb realize)
  - (d) The modifiers could attach to the N cheese, either one or both – the cheese is found in the bathtub or that Jerry eats the cheese that is present in the bathtub on Thursday

## Word order differences

Different word-order patterns can be generated by changing the relative order of specifiers and complements with respect to the head.

Example:	
Japanese: 'Head-final' language	Thai: 'Head-initial' language
Tokyo kara	sii nii
Tokyo from	colour this
Masaga ringoo tabeta	X' → (Mod) Comp X
Masa apple ate	Xp → Spec X'

Daikoga Ayumiga Kazukoo horneta to itta (like Hindi)  
[Daiko ([Ayumi Kazuko praised] that) said]  
Daiko said that Ayumi praised Kazuko.

## Phrase structure - summary

- Sentences have complex structure, built out of many simpler pieces
- Given the rules, grammatical sentences can be built, ungrammatical ones cannot
- Accounts for important aspects of linguistic creativity – finite means infinite output
- Structural ambiguity can be explained
- Cross-linguistic variation in the linear order of elements are easily accounted for

### Other Word Order Patterns

- S → NP VP  
VP → V NP  
yields "SVO" word order pattern
- Which of the following word orders can be generated by reordering elements of these two rules?  
SVO, SOV, VOS, OVS  
[accounts for rarity of OSV]
  - Two common word orders that we can't yet account for: VSO (Irish), verb-second (German, Kashmiri)
  - In order to explain how to generate these word orders, we need to take a diversion into an area of English syntax that we also can't account for with our current set of rules.

## Transformations

English has yes-no questions formed by displacing the auxiliary:  
 Will **the cat** chase the dog?  
 Can **the dog** flee the cat?

The auxiliary appears to the left of the subject rather than before the verb with the associated intonation change. The PSRs cannot generate Aux before subjects:  
 $TP \rightarrow Spec\ T'$   
 $T' \rightarrow TVP$  (where T contains the auxiliary element bearing tense)

We say the trees generated by PSRs may be modified in that contents of certain branches can be moved to other positions in the tree. Transformations are powerful devices so we must provide good reasons for including them in our grammars and be careful in how we use them.

## Where are they moving to?

Aux inversion is actually, T to C movement  
 Wh-movement is movement to Spec, CP  
 These are the two locations that are made available by positing a CP-shell structure that lies above all sentences (TPs)

Note:

- A transformation only moves a constituent, it cannot change categories, it cannot eliminate any part of the structure that has been created by the PSRs
- The position vacated by the moved element is not empty/open/free. It is filled by an empty element called *trace/ei* which is co-indexed with the moved constituent

## Advantage transformation

- (1) We don't have to make a disjunctive statement about auxiliary verbs: that they sometimes appear before verbs and sometimes before subjects. Rather, we can say that all auxiliaries are base-generated in the same position (T). The inversion transformation then applies and allows us to form questions of a certain type.
- (2) The account where a base-generated tree is altered in some way, captures speakers' intuition that *Will the cat chase the dog?* And *Can the dog flee the cat?* correspond to the statements *The cat will chase the dog* and *The dog can flee the cat*.

## Why T to C?

Returning to embedded clausal complements, we said that clause introducers like *if*, *whether* and *that* are Complementizers that project their own phrase, CP:  
 The coach wondered [<sub>CP</sub> if/whether the student should play]  
 A fan asked [<sub>CP</sub> if the team will win]  
 If we attempt yes-no questions within the embedded clause we find  
 \*The coach wondered [<sub>CP</sub> if/whether should the student t play]?  
 \*A fan asked [<sub>CP</sub> if will the team t win]?  
 If the complementizer is absent, yes-no questions are better  
 The coach wondered [<sub>CP</sub> should the student t play]?  
 A fan asked [<sub>CP</sub> will the team t win]?

## Wh-movement – another question

Consider the following sentences  
 Which car should the man fix?  
 What can the child sit on?  
 Here there are two constituents in unexpected places: (a) the wh-phrase and (b) the auxiliary. The base sentences are:  
 The man should fix **which car**. (object of verb)  
 The child can sit on **what**. (object of preposition)  
 It is not possible to either say  
 \*The man should fix (too few arguments)  
 \*Which car should the man fix the scooter (too many arguments)

## Other features of T-C movement

- T can move across both pronominal subjects and fully referential NP subjects:  
 Will **he** t say the prayer?  
 Will **the priest** say the prayer  
 Will **the archbishop of Canterbury** t say the prayer?
- Only auxiliaries can move to C not main verbs, in Modern English  
 Will he pray?  
 \*Pray he?  
 • However, in Elizabethan English one could say  
 Saw you my master?  
 Speakest thou in sober meanings?  
 Know you not the cause?  
 Spake you not these words plain?

## Relative clauses, wh-movement revisited

- Sue knows the man [who Bob criticized]
- *See handout on extending wh-movement*

## Sentence reconsidered

- The relation between subject and predicate is indicated through agreement (also called Spec-Head agreement)
- Information regarding the time and aspect of the event indicated by the verb are not confined to the verb but range over the whole sentence.
- Let's posit a node T that builds TPs which replace our Ss. Rewriting the PSR for S and CP we have:  
 $TP \rightarrow (\text{Spec}) T (\text{Comp})$ , Spec is NP and Comp is VP  
 $CP \rightarrow (\text{Spec}) C (\text{Comp})$ , where Comp is TP

## THE PRINCIPLES OF FIRST LANGUAGE ACQUISITION

WHAT FIRST LANGUAGE ACQUISITION TELLS US ABOUT THE COGNITIVE SYSTEM OF LANGUAGE

1

The acquisition of language "is doubtless the greatest intellectual feat any one of us is required to perform".

--Leonard Bloomfield, *Language*, 1933

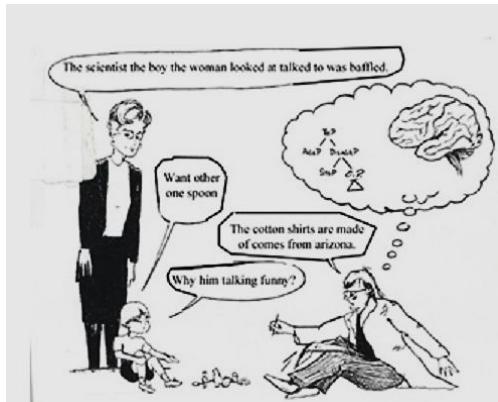
The capacity to learn language is deeply ingrained in us as a species, just as the capacity to walk, to grasp objects, to recognize faces. We don't find any serious differences in children growing up in congested urban slums, in isolated mountain villages, or in privileged suburban villas.

--Dan Slobin, *Human Language Series 2*, 1994

How comes it that human beings, whose contacts with the world are brief and personal and limited, are able to know as much as they do know.

--Bertrand Russell, *Human Knowledge: Its Scope and Limits*, 1948

4



## DO YOU REMEMBER LEARNING TO...

tie your shoe laces?

ride a bicycle?

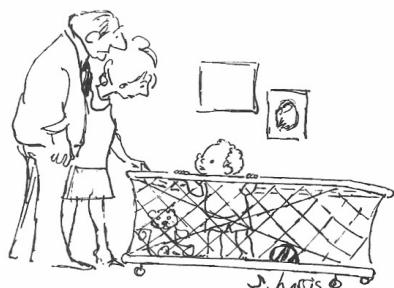
to count?

to speak?

You may recall the first three. But not the fourth – other than anecdotes of baby talk from parents

And you received explicit instruction and training to perform the first three but not to speak your first language(s)

5



"WHAT'S THE BIG SURPRISE? ALL THE LATEST THEORIES OF LINGUISTICS SAY WE'RE BORN WITH THE INNATE CAPACITY FOR GENERATING SENTENCES."

Copyright © S. Harris.

3

## HOW IS THIS POSSIBLE?

What does the child do or not do in acquiring a language?

Children do not and cannot store all possible words and sentences since children learn to understand and produce (an infinite number) of novel sentences.

Children learn the "rules" to use their language creatively.

Children learn the words, sound and sentence structure of their native language(s).

6

## LANGUAGE LEARNING IS AN INNATE CAPACITY

Characteristics of biologically controlled behaviours:

1. The behaviour emerges before it is necessary
2. Its appearance is not a result of conscious decision
3. Its emergence is not triggered by external events (though the surrounding environment must be sufficiently *rich* for adequate development)
4. Direct teaching and intensive practice have relatively little effect
5. There is a regular sequence of *milestones* as the behaviour develops and these can usually be correlated with age and other aspects of development
6. There is likely to be a critical period for the acquisition of the behaviour

7

## PHONETICS AND PHONOLOGY

Some techniques:

**High Amplitude Sucking** with children upto six months of age (sucking produces a noise that they hear – some phone)

**Conditioned Head Turn Procedure** for infants aged 5-18 months

**Heart beat (prenatal responses)**

**Preferential looking (up to 2-3 years)**

10

## STAGES OF LANGUAGE ACQUISITION-1

Language acquisition follows predictable sequences across languages and cultures but the absolute ages varies from child to child. The acquisition of abilities is also gradual .

Stage	Age	Description
Comfort	0;2-0;4	no movement of vocal tract
Vocal plays	0;4-0;6	squeals, growls, snorts, gurgles
Babbling	0;6 - 0;8	repetitive CV patterns
One-word	0;9 - 1;6	single words or word stems
Two-word	1;6 - 2;0	mini-sentences
Multiword 1	2;0 - 2;6	telegraphic" sentences with <i>lexical</i> not <i>functional</i> morphemes
Multiword 2	2;6 on	<i>functional</i> morphemes

This inventory appears to be universal, perhaps for bio-mechanical reasons?

8

## IN THE WOMB

**De Caspar and Spence (1986)**

Mothers read aloud Dr. Suess' *Cat in the Hat* in the last six months of pregnancy

After birth children heard many stories including *Cat in the Hat*

Infants who had heard the story modified their sucking rate; infants who hadn't heard the story did not show any change

11

## STAGES OF LANGUAGE ACQUISITION-2

What do children know at various stages of development? Is their production a direct representation of their perception/knowledge?

What is innate and what is learned? We discuss examples from the grammatical modules of

- Phonetics and Phonology
- Lexicon, word learning
- Morphology
- Syntax

9

## SOUND-VISUAL CUES CORRELATION

Differentiating between [a] and [i]

Video of two faces saying [a] and [i] and audio playing one of the two sounds

Head turn to match audio cue, matching the visual cues

Children own cooing noises match the audio-video they prefer

12

## PHONOLOGY: CATEGORIZATION - 1

Phoneme categorization (Eimas, Siqueland, Jusczyk, and Vigorito 1971):  
Issue: Do infants perceive speech sounds categorically?

Subjects: 1 and 4-month-old babies acquiring English

Stimuli: tokens on the /ba-/pa/ continuum (<25ms /b/)

VOT (msec)= 0 20 40 60 80

Adult perception: [-----ba-----] [-----pa-----]

Babies sucked on pacifiers connected to a computer. First one syllable was played over and over, then the stimulus changed.

Pairs of input: [20, 40], [60, 80], [0,20]

Results: increase no change no change

Reason: Category change for first pair, novel stimulus

13

## PHONOLOGY: CATEGORIZATION - 4

The effect of language experience on perception (Werker and Tees 1984)

6-8 month-old infants discriminate between contrasts not made in English (retroflexes in Hindi) like Hindi speaking adults do but not English-speaking adults; 95% accuracy

8-10 months decline in discrimination ability, 60-70% accuracy

10-12 months apparent loss of ability, 20% accuracy

16

## PHONOLOGY: CATEGORIZATION - 2

Are these categories learned or innate? Lasky, Syrdal-Lasky, and Klein (1975)

Subjects: 4-6.5-mo old Guatemalan babies acquiring Spanish

Stimuli: tokens on the /ba-/pa/ continuum

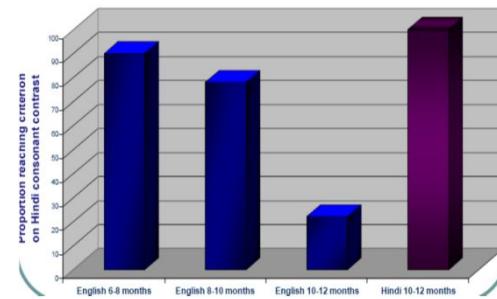
[20, 60] voiced-voiceless boundary for English speakers

[-20, 20], boundary between voiced-voiceless stops of Spanish

[-60, -20] boundary between Thai pre-voiced and voiced stops

Do the infants' responses change because Spanish VOT is different from English VOT? Are they sensitive to a categorization (prevoicing) that is not seen in their language?

14



17

## PHONOLOGY: CATEGORIZATION - 3

VOT -60 -20 20 60

Spanish adults: [-----ba-----][-----pa-----]

English adults: [-----ba-----][--pa--]

Thai adults: [-bba-][-----ba-----][--pa--]

Results:

20/60 Voiced-voiceless English stops Increase in heartbeat

-60, -20 Prevoiced-voiced Thai stops Increase in heartbeat

(Same results found with Kikuyu and English speaking children)

-20, 20 Voiced-voiceless Spanish stops NO CHANGE

15

## BABIES ARE UNIVERSAL LISTENERS

Kikuyu infants can distinguish /ba/ vs. /pa/, a contrast not present in Kikuyu (Kenya)

Canadian (English) infants can pick out non-English contrasts in Polish and Czech

Japanese/Korean infants can distinguish /r/ vs. /l/ (but their parents can't!)

18

## IS CATEGORICAL PERCEPTION UNIQUE

To language? No. Colour perception also shows categories  
 To humans? No. Chinchillas, quails, etc. can distinguish the /b/, /p/ categories as well  
 Perceptual continuity with other species but co-opting this perception for speech and in how we organize consonants and vowels is unique to human beings

19

## BABBLING

Deaf children and hearing children babble manually and orally  
 Oral babbling is limited in deaf children and falls off very soon, as is the gestural babble of hearing babies.  
 Use a reduced inventory of SL gestures  
 repetitive  
 meaningless  
 has continuity in the first usage of words (i.e., gestures)

22

## PHONOLOGY - PRODUCTION

Production lags behind perception: 'Fis' phenomenon  
 (Berko and Brown 1960)

Child (pointing at aquarium): /fis/  
 Adult Huh?  
 Child /fis/  
 Adult Oh, fis.  
 Child No fis  
 Adult Oh! Fish!  
 Child Ye fis

20

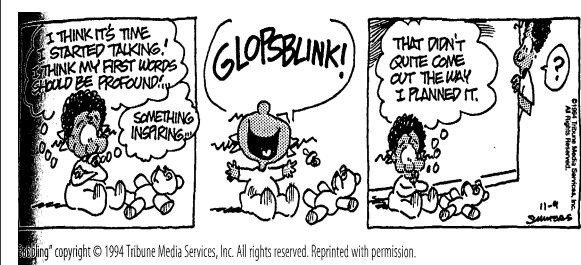
Much babbling happens when the child is alone – practice for future use, gaining muscle control  
 Babbling is maturationally driven – children babble at the same ages  
 Evidence? Children with tracheotomy – when tube removed begin to babble. Babbling signals brains maturity  
 Pitch variations are also seen

23

## PHONOLOGY - PRODUCTION

Baby tongue is large relative to the size of the oral cavity – palatals and bilabials  
 Even babbling is not linguistic chaos – the 12 most frequent consonants in the world's languages make up 95% of the consonants that the infants use while babbling.  
 Initial babbles are repeated CVCV sequences – repeated or canonical babbling.  
 Gradually babbling uses target phonemes of the language  
 Variegated babbling – mixed syllables  
 Order of manner and place acquisition:  
 >>Nasals, glides, stops, liquids, fricatives and affricates; labials, velars, alveolars and palatals

21



"Babbling" copyright © 1994 Tribune Media Services, Inc. All rights reserved. Reprinted with permission.

24

## PHONOLOGY

Mapping sounds and meaning: much mispronunciation and misphonology

Children typically utter their first words toward end of the first year.

There is lots of variation in the types of words uttered.

Early words consist of 1 or 2 syllables

General increase in use of true consonants

- labials more frequent than coronals
- coronals more frequent than velars
- stops, nasals more common, fricatives less frequent – they require more precise articulatory control

25

## INVENTORIES

### Word Initial Position    Word Final Position

m- n-	-n
b- d- g-	-p -t -k
p- t- k-	-s
f- s- h-	
w- j-	

Variation between children, across languages

26

Vowels tend to be the basic “cardinal” vowels [i a u]

CV is the preferred syllable structure

Omission of word final consonant

Omission of fricative or liquid in cluster

26

## PHONOLOGICAL PROCESSES

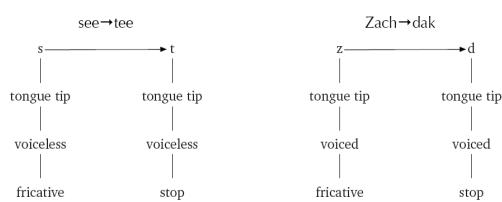
Daniel (Menn 1998)

<u>Set 1</u>	<u>Set 2</u>
bump	gug (for "bug")
down	guck (for "duck")
gone	gig (for "pig")

Why? It is not that he cannot say /b/, /d/ and /g/

29

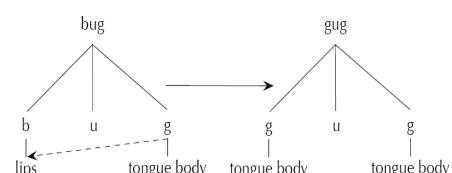
## FEATURES



Some features are retained and some are not

27

## ASSIMILATION OR SPREADING



maliki-kirli-rlı-lki-ji-li ‘the dogs are with me now’  
kurdu-kurlu-rlı-lku-ju-lu ‘the children are with me now’  
(Warlpiri)

30

## SYLLABLE SHAPES

First words are simple – segment wise and syllable shape wise

Stoel-Gammon results on 33 2-yr-olds use of syllable shapes

CV, V (all samples)

CVC (97%)

CVCV (79%)

CVCVC (65%)

31

## CONSONANT SIMPLIFICATIONS

Use a simpler consonant for a more difficult one

**Fronting:** consonants produced at the forward part of the mouth are easier than those produced further back, 'd' & 't' easier than 'g' or 'k', 'cow-dow'

**Stopping:** sound that stop airflow are easier than ones that impede it; 't', 'd' are easier than 's', 'z', 'sh'; 'p', 'b' are easier than 'f', 'v', foot-put

**Gliding:** replace 'r' and 'l' with 'w'; e.g. rock-wo(ck)

34

## MODIFICATIONS

omission:

water-wa, bottle-ba, banana -nana, Tyrannosaurus Rex-Toy Rex

(keep it CV like)

32

## 12 MONTHS AND BEYOND

### Holophrastic stage 12-24 months

- idiomorphs
- mutual exclusivity and whole object bias
- overgeneralizations, undergeneralizations
- referential vs. expressive children

### Two-word Stage—24-30 months

- Subject-verb 'Mary go.'
- Verb-modifier 'Push truck.'
- Possessor-possesed 'Mommy sock'
- Content words, no function words

### Telegraphic Stage—30-36 months

- 2-5 words with little extra morphology
- Morphological overgeneralization
- Easier, more productive morphemes first

35

## SYLLABLE PROCESSES

### Final Consonant Deletion:

omit word-final consonants to preserve CV; back-baa, cat-ka, soap-so

### Unstressed Syllable Deletion: delete syllable with weak stress

e.g. banana-nana, balloon-loon, telephone-tephone, dessert-zert, helicopter-koter (note that stressed syllables are generally preserved; Adults: lab, rhino, dorm, fax, cell etc.)

### Cluster Reduction:

delete consonant combinations, e.g. 'play' spoken as 'pay'

### Addition:

duck-ducke, car-kaka, piano-pinano

33

## MLU

Stage 1 - 12 to 26 months =

MLU 1.00 to 2.00

Stage 2 - 27 to 30 months =

MLU 2.00 to 2.50

Stage 3 - 31 to 34 months =

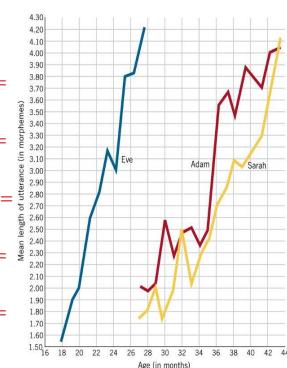
MLU 2.50 to 3.00

Stage 4 - 35 to 40 months =

MLU 3.00 to 3.75

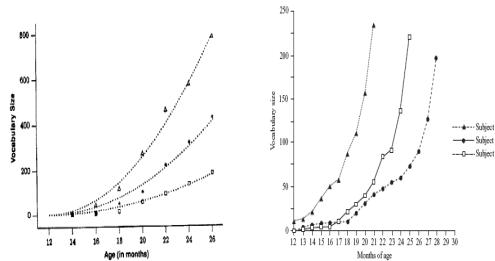
Stage 5 - 41 to 46 months =

MLU 3.75 to 4.50



36

## VOCABULARY GROWTH



37

## NOMINALS

Specific: individuals

General: all members of a category

Problem: how to distinguish the two

e.g. 'daddy' (all men)

'doggie' (family dog)

### Categories of Nominals

Sound Effects, 'meow', 'moo'

Food and Drink, 'apple', 'juice'

Body Parts & Clothing, 'eye', 'hat'

House & Outdoors, 'cup', 'tree'

People, 'baby', 'daddy', 'mommy'

Toys & Vehicles, 'ball', 'car'

40

Italian vocabulary learning seems to be slower than English, ditto for Icelandic; no difference for Hebrew, Spanish, Finnish

Female advantage in English, but not in Swedish

Mother's education level is positively correlated: birth order and father's education level data inconsistent

Word spurt: some children spurt earlier than others, perhaps not all children spurt

English/French has noun dominance, but Korean/Chinese has verb dominance

38

## OTHERS

### Action Words

General Actions: 'down', 'go', 'eat'

Locatives: 'in', 'on', 'table'

Social-Action Games, 'peekaboo'

### Personal-social: Affective states and relations

Assertions: 'yes', 'no', 'want'

Social-Expressive Actions, 'hi', 'byebye', 'nite-nite'

### Modifiers: Properties & qualities of things or events

Attributes, 'big'

States, 'allgone', 'hot'

Locatives, 'there'

Possessives, 'mine'

41

Benedict (1979)

- 8 Children
- Longitudinal Study (Six Months)
- Comprehension (C) and Production (P)
- Rate and Gap between C & P

Age	Comprehension	Production
0;10	0	
0;11	20	
1;1	30	0
1;3	50	20
1;4		30
1;5		40
1;6		50

39

## TWO KINDS OF CHILDREN

Expressive: Higher use of social-personal words

Referential: Dominance of general nominals

Categories	Referential	Expressive
Specific Nom.	6%	14%
General Nom.	76%	34%
Action	8%	12%
Personal-social	2%	24%
Modifiers	4%	12%
Other	2%	4%

42

## LEARNING WORD MEANING

Innate biases or expectations that constrain the hypothesis space for learning word meaning

**Whole object bias:** Nouns must refer to whole objects.

**Taxonomic bias:** Nouns refer to classes of objects and not to one specific example. But children only hear specific examples!

**Shape bias:** shape plays a more important role in defining object categories than color, size, material, etc.

**Mutual exclusivity bias:** words always have different meanings (ranked above Whole object bias)

43

- Look at the merf! It's a merf! Look, a merf!

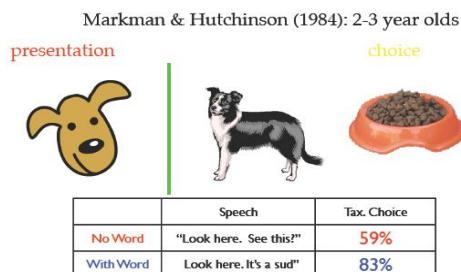


- Look at this one! Isn't it neat? It's cool!



- Please hand me the wug.

46



44

## ACQUIRING MORPHOLOGICAL RULES

This was the first evidence that psycholinguists found that kids are developing a grammar; they can produce new words that they never heard before. (Jean Berko 1958)

Test: "Here is a wug. Now there are two of them. There are two ...?" (Answer, /wugz/)

(English plural rule, with allomorphic variation)

heafs, wugs, luns, tors, cras, tasses, gutches, kashes, nizzes

Results: Three year olds perform with great accuracy

47

Words refer to objects of like kind, not like "themes" (e.g., relatedness: Markman & Hutchinson 1984)

So, "dog" refers to dogs, but not dog collars, dog owners, dog food etc.

This is not obvious since children under 6 sort thematically without verbal cues

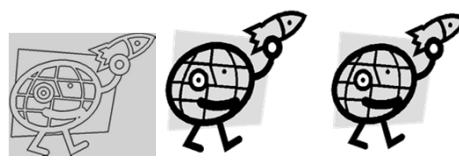
So naming does something different to the world view

45

## WUG TEST

This is a wug

Now there are two of them. There are two....



48

## ACQUIRING MORPHOLOGICAL RULES

English past tense morphemes:

My teacher holded the baby rabbits and we patted them  
 Hey, Horton heared a Who  
 I finded the ball  
 The alligator goed kerplunk.  
 Not produced by adults, over-regularization.

49

## IRREGULARS

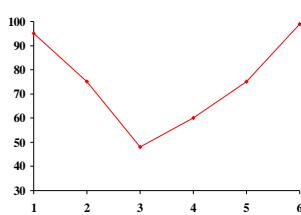
1. went	men	worst
2. goed	mans	baddest
3. wented	mens	workest
4. went	men	worst

52



This is a bird who knows how to rick. It is ricking.  
 What did it do yesterday? Yesterday it...

50



53



This is a frog who knows how to mot. He is motting right now.  
 What did he do yesterday? Yesterday he ....

51



"Baby Blues" copyright © 1997 Baby Blues Partnership. Reprinted with special permission of King Features Syndicate, Inc.

54

## WHAT WE LEARN FROM THE WUGS

The major finding of the wug test was that even very young children have already internalized systematic aspects of the linguistic system that enable them to produce plurals, past tenses, possessives, and other forms of words that they have never heard before.

The test has been replicated many times, and it has proven to be very robust.

It was the first experimental proof that young children have extracted generalizable rules from the language around them.

55

## ACQUIRING MORPHOLOGICAL RULES

Gordon 1986

Subjects: 3-5 year olds

"Here is a monster that eats mud. He is called a mud-eater."

"... who likes to eat mice" MICE-EATER (MOUSE-EATER)

"... who likes to eat rats" RAT-EATER

"... who likes to eat purple people" PURPLE-PeOPLE-EATER

"... who likes to eat purple lions" PURPLE-LION-EATER

Question: Learning from parents by looking at plural-compounds?

No plural compounds in motherese (poverty of stimulus)

Automatically distinguish between stored and derived items.

58

This is a little wug. What would you call such a small wug?

This wug lives in a house. What would you call a house that a wug lives in?



56

## FUNCTIONAL CATEGORIES AND LEXICAL CATEGORIES

Children begin to acquire functional categories around the time they begin to make word combinations

They are acquired over several years

59

## ACQUIRING MORPHOLOGICAL RULES

**Compound formation:** Kiparsky noted that compounds can be formed from irregular plurals but not regular plurals:

- |                 |                |
|-----------------|----------------|
| • men-bashing   | *gays-bashing  |
| • mice-infested | *rats-infested |
| • teethmarks    | *clawmarks     |

Irregular plurals are stored in the mental dictionary as roots or stems and can enter word-formation; compounding takes place with stored words.

57

## FACTORS

Semantic Complexity, e.g. {-ing} vs. {-ed}

Syntactic Complexity, e.g. agreement present tense {-s} vs. past tense {-ed}

- 'I walk' 'he walk/s'
- 'I walk/ed' 'he walk/ed'

Perceptual Salience, e.g. present {-s} vs. progressive {-ing}

Frequency,

- e.g. {-s} vs {-ing}
- e.g. 'in' vs. 'between'

60

### Allomorphy: variants of a single morpheme

- E.g. 'plural' {-s}
- cats /s/ dog /z/ bush /ez/

- children need to acquire all the variants

{-ing} has no variants

Generality of rule application

61

### ACQUIRING SYNTAX

(Hirsh-Pasek et al. 1988)

Found the same result for 24-month-olds with more subtle syntactic cues, and without any lexical cues:

"Big Bird is flexing Cookie Monster." [CAUSE TO]

"Big Bird is flexing with Cookie Monster." [ALONG WITH]

Thus young kids can use the syntactic cues in comprehension even before they can produce whole sentences.

64

### ACQUIRING SYNTAX

Do children have syntax in the one-word stage? (Hirsh-Pasek et al. 1985)

Subjects: 17-mo-olds, one-word stage, using few/no verbs.

They watched two TV's showing familiar characters from the Sesame Street TV show: Cookie Monster and Big Bird.

Task: One TV showed CM tickling BB; the other TV showed BB tickling CM. The babies heard one of these sentences:

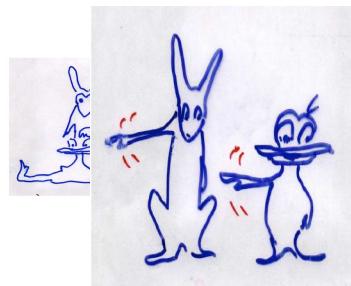
(1) "Cookie Monster is tickling Big Bird."

(2) "Big Bird is tickling Cookie Monster."

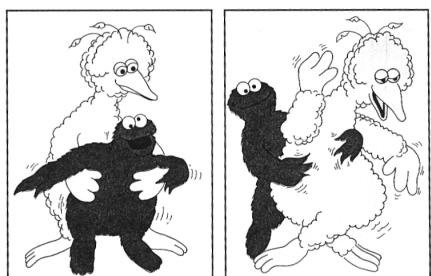
Results: the babies looked longer at the correct TV scene

62

### GORP – RABBIT/DUCK



65



63

### ACQUIRING SYNTAX

Children make certain errors (goed, holded, oxes etc.) but never certain others that are logically plausible

In all languages it is impossible to move a "wh-phrase" out of a conjoined noun phrase (as in 2b)

1a. John ate eggs with toast

1b. What did John eat eggs with?

2a. John ate eggs and toast?

2b. \*What did John eat eggs and?

66

## ACQUIRING SYNTAX

### Knowledge of wanna-contraction in English

- 1a I want to eat a cookie.
- 1b I wanna eat a cookie.
- 2a What do you want to eat?
- 2b What do you wanna eat?
- 3a You want who to eat a cookie?
- 3b Who do you want to eat a cookie?
- 3c \*Who do you wanna eat a cookie.

**Generalization:** contraction is not possible when questioning the subject of the subordinate clause.

67

## STORY



Hello, Eeyore! I see that you are reading a book.

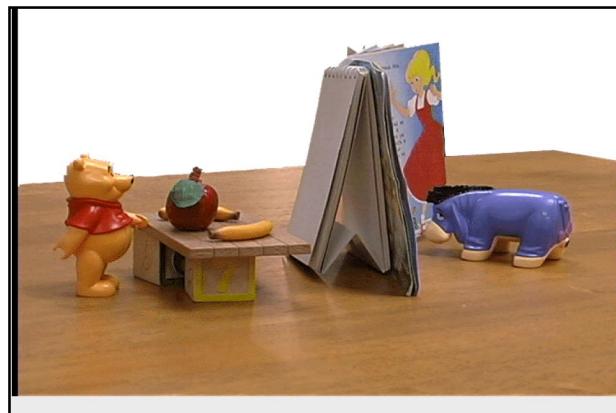
70

## ACQUIRING SYNTAX

An Experiment to test *wanna*-contraction in children aged 3-5 years. The experimental protocol is designed to elicit questions in which the *object* of the subordinate clause is questioned

- Exp: The rat looks hungry. I bet he wants to eat something. Ask Ratty what he wants.  
 Child: What do you wanna eat?  
 Rat: Some cheese would be good.

68



What a fine-looking apple!

## ACQUIRING SYNTAX

The next protocol is used to elicit questions about the *subject* of the subordinate clause

Exp: There are three guys in this story: Cookie Monster, a dog, and this baby. One of them gets to take a walk, one gets to take a nap, and one gets to eat a cookie. And the rat gets to choose who does each thing. So, *one* gets to take a walk, right? Ask Ratty who he wants.

- Child: Who do you want (\*wanna) to take a walk?  
 Rat: I want the dog to take a walk.

69

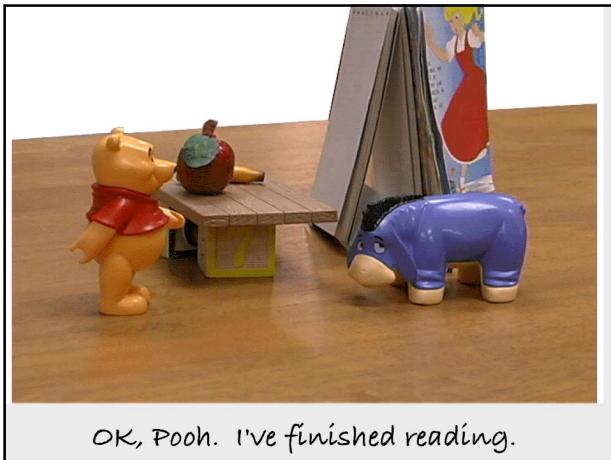




OK, I'll have to eat a banana instead.



I shouldn't be such a greedy donkey --  
I should let Pooh eat this apple.



OK, Pooh. I've finished reading.



I suppose I can eat a banana instead.



Great. Now that Pooh is reading the book,



Here you are, Pooh. You can have the apple.



Oh, I'm such a lucky bear! I can read the book and I can eat the apple, at the same time.

### How 3-4 Year Olds Did

- While Pooh was reading the book, he ate an apple.
- While he was reading the book, Pooh ate an apple.
- Pooh ate an apple while he was reading the book.
- He ate an apple while Pooh was reading the book.

yes!  
yes!  
yes!  
no!

Works for English, Italian, Russian, etc.



Apple is eaten up.

### LEARNING IN BEHAVIOURIST TERMS

B.F. Skinner (1904-1990) claimed that language is just another form of behavior. It is a response to stimuli in the environment. And it is learned.

Predictions of behaviorism for first language acquisition:

- Children would make many errors.
- Errors would differ from child to child.
- The level of linguistic attainment would differ from child to child.

83



OK. That was a story about Eeyore and Pooh. First Eeyore was reading the book then Winnie-the-Pooh was reading the book. I know one thing that happened.

While Pooh was reading the book, he ate the apple.  
While he was reading the book, Pooh ate the apple.

He ate the apple while Pooh was reading the book.

### LEARNING AS CONDITIONING

Ivan Pavlov (1849-1936) discovered Classical Conditioning. Learning is the involuntary association of stimulus and response.



84

## CLASSICAL CONDITIONING

**Neutral stimulus:** Bell ring  
**Unconditioned stimulus:** Present food  
**Unconditioned response:** The dog salivates  
**Conditioning:** Repeat the bell ringing and food presentation several times.  
**After conditioning:** The dog hears a bell and then the food

**Pavlov's Classical Conditioning:** Learning is the association of stimulus and response.

- Neutral stimulus: Bell rings
- Unconditioned stimulus: Present food
- Unconditioned response: The dog salivates

**After conditioning:** The dog hears a bell and then the dog salivates even when no food is presented.

85

## OPERANT CONDITIONING IN LANGUAGE

**Shaping** is the process of teaching a complex behavior by rewarding closer and closer approximations to the desired behavior.

Reinforce any behavior that is similar to the behavior you want.

Reinforce only those behaviors that are close to the desired behavior.

Reinforce only the desired response.

**Behaviorist Learning of Language?**

Adults train the child to say "water" when it sees water.

Reinforce "wa", "wawa", etc.

Reinforce "wada"

Reinforce only "water"

88

## CLASSICAL CONDITIONING IN LANGUAGE



**Adult utters "water" whenever the child sees water.**

**Unconditioned stimulus:** "water"  
**Unconditioned response:** the child . . . Drinks?  
 Splashes? Spits?

**After conditioning:**  
**Conditioned stimulus:** the word "water" –  
**Conditioned response:** The child drinks.

86

## BEHAVIOURIST THEORIES OF ACQUISITION - IMITATION

**Children produce what they hear.**

Child: My teacher holded the baby rabbit and we patted them.

Adult: Did you say that your teacher held the baby rabbit?

Child: Yes

Adult: What did you say she did?

Child: She holded the baby rabbit and we patted them

Adult: Did you say she held them tightly?

Child: No, she holded them loosely.

89

## OPERANT CONDITIONING



**Learning in which a voluntary response is strengthened or weakened, depending on its positive or negative consequences**

**Works with voluntary muscles only, in contrast to Pavlov's classical conditioning.**

A hungry pigeon is placed in a box. When, by chance, the pigeon pecks a key in the box, it receives food.

The first time this happens the pigeon does not learn the connection between pecking and receiving food.

Eventually, the pigeon pecks the key continuously until its hunger is satisfied.

The pigeon demonstrates that it has learned the connection between stimulus and response.

87

## IMITATION THEORY: PROBLEMS

**Non-imitation of parents**

**Systematic errors across children and languages:**

hitted, goed (overgeneralization)  
 no drink (He does not want a drink)  
 dog toy (That's the dog's toy)

**Producing novel sentences**

**PLD is full of questions and imperatives, children produce declaratives mainly**

**Children who are unable to speak for any neuro-physiological reason, learn the language spoken to them and understand it. They use the language for speaking once they overcome their impairment.**

90



91

## PROBLEMS WITH EXPLICIT INSTRUCTION/REINFORCEMENT

**Parental reinforcement /instruction is seldom seen**

**Parents correct for meaning, not form (Brown & Hanlon 1969)**

**Child: Mama isn't boy, he a girl.**

Adult: That's right

Child: Walt Disney visits on Tuesday.

Adult: No, he doesn't.

**Much of adult's knowledge is NOT explicit**

**When parents do correct, children don't get it**

94

## REINFORCEMENT, EXPLICIT INSTRUCTION

**Reinforcement Theory: Children learn through explicit teaching with positive/negative reinforcement**

Child: Want other one spoon, Daddy.  
 Daddy: You mean, you want THE OTHER SPOON?  
 Child: Yes, I want other one spoon, please, Daddy  
 Daddy: Can you say 'the other spoon'?  
 Child: Other...one...spoon.  
 Daddy: Say 'other'  
 Child: Other  
 Daddy: 'Spoon'  
 Child: Spoon  
 Daddy: 'Other...Spoon'  
 Child: Other...spoon. Now give me other one spoon?

92

## LEARNING BY ANALOGY

**Analogy Theory: Children learn language by hearing a sentence and creating analogous sentences**

**Problems:**

I painted a red barn. I painted a blue barn.

I painted a barn red. I painted a barn blue

I saw a red barn.

I saw a barn red (not possible, not produced)

sun → sunny

moon → moony (not possible)

95

## MORE INSTRUCTION...

Child: Nobody don't like me  
 Adult: No say. "Nobody likes me."  
 Child: Nobody don't like me  
 (dialogue repeated eight times)  
 Adult: Now, listen carefully, say "Nobody likes me."  
 Child: Oh, nobody don't likes me.

93

## LEARNING THROUGH STRUCTURED INPUT

**Children learn because adults talk to them in a special way – simplified language – called Motherese or baby talk or Infant Directed Speech**

**Infants do show auditory preference for motherese with its exaggerated intonation contours**

**However,**

**Motherese is not simple, it contains complex syntactic structures  
 Not all cultures have motherese, but all children end up speaking anyway**

96

## SO...

Language is not really something that the child does; it is something that happens to the child placed in an appropriate environment, much as the child's body grows and matures in a predetermined way when provided with appropriate nutrition and environmental stimulation.

Noam Chomsky, *Managua Lectures*

97

## THE CRITICAL PERIOD HYPOTHESIS

The difference between adults and children is in the ease of language learning – the effect of 'accent', 'grammatical mistakes' etc. are attributed to the *critical period*

Acquisition by the age of 7 yields native command

Acquisition between the ages of 8 and 15 yields progressively less perfect command

The innate ability to acquire language degrades with age.  
Suggests that kids are "built to learn language" in a way that adults are not.

100

## INNATENESS HYPOTHESIS

The solution to the logical problem of language acquisition – accounting for the ease, rapidity and uniformity of acquisition in the face of impoverished data.

Humans are born 'ready' for language – Universal Grammar is hardwired.  
Acquisition is rapid and completed by age four-five  
Poverty of the stimulus – lack of negative evidence and sometimes positive evidence  
Input may be fallible: speech errors, false starts, ungrammatical and incomplete sentences  
Exposure combined with UG leads children to construct a grammar

98

## LENNEBERG 1967

Lenneberg 1967 (or Penfield and Roberts 1959) is usually considered to be the written origin of this idea that there is a "critical period" or "sensitive period" for language acquisition.

He based this on several observations, including the observation that critical periods are biologically common

101

## EVIDENCE FOR INNATENESS/LAD

**Universals in acquisition**

**Children neither imitate nor are they instructed**

**Innate perceptual categories**

**'Invention' of language by children deprived of real input**

**Critical age – normal cognition but abnormal language**

**Abnormal cognition and normal language (William's syndrome, hydrocephalics)**

**Inherited language deficiencies (special language impairment)**

99

## EVIDENCE FOR CPH

Genie, isolated until the age of 13.5, never learned to produce more than telegraphic speech.

- Applesauce buy store.
- Neal come happy: Neal not come sad.

Another case of severely delayed language access (but without abuse) is Chelsea, misdiagnosed as retarded in early childhood, when in fact she was congenitally deaf—only discovered when Chelsea was 31. Chelsea's utterances have almost no discernible structure at all; her speech was less language-like than Genie's.

Isabelle, isolated until the age of 6.5, mastered grammar within one year

- Why does the paste come out if one upsets the jar?
- Do you go to Miss Mason's school at the university?

Jim, hearing child of deaf parents, brought into speech contact around 3;6, rapidly caught up in spoken language, reaching age-norms by 6.

102

## CPH IS NOT RESTRICTED TO LANGUAGE

Aspects of maturation are seen in humans and animals.  
in ducklings: imprinting, ability to identify and follow the mother  
in kittens: ability to perceive visual images  
in sparrows: ability to learn the father's songs  
Developing the neural circuits for such skills is an expensive allocation of developmental resources and evolution favours individuals who lose this once learning has (normally) occurred  
The visual system is a favourite example. In experiments done on macaque monkeys, it was determined that there is a critical period for development of binocular vision cells in the visual cortex (tested by monocular deprivation)  
In fact, vision studies "replicated" in humans as well; there seems to be a visual critical period at around age 6, after which providing previously delayed visual stimuli is of no use. (Congenital opacities of the cornea; surgery performed on juveniles or adults does not restore sight)

103

## CHILD APHASIA AND CPH

**Acquired aphasia during childhood is almost never fluent (mutism), but they recover rapidly (lasting effects, generally only slight word-finding and vocabulary difficulties)**  
**Recovery is faster, better than in adult acquired aphasia, but not complete.**  
**Early enough, right hemisphere can take over language functions after a serious loss in the left hemisphere, but it doesn't do as good a job (reversible sentences).**

106

## NEUROSCIENTISTS ON CPH

...The development of form perception and the binocular vision necessary for depth perception proceed in stages after birth. Each stage culminates in one or more developmental decisions, many of which are irreversible. In each stage, appropriate sensory experiences are necessary to validate, shape, and update normal developmental processes. Consequently, the effects of sensory deprivation are most severe during a restricted and well-defined period early in postnatal life when these developmental decisions are still being made. (*Kandel, Schwartz, Jessell* 3d ed. 1991, p. 956)

104

## CHILD APHASIA

**Lenneberg's summary of the results of left hemisphere lesions as a function of age:**

- 0-3 months: no effect
- 21-36 months: all language accomplishments disappear; language is re-acquired with repetition of all stages.
- 3-10 years: aphasic symptoms, tendency for full recovery
- 11th year on: aphasic symptoms persist.

**Basis for his view that lateralization was tied to critical period.**

107

Critical periods of development generally do not have sharp time boundaries. Different layers within one region of the brain may have different critical periods of development, so that even after the critical period for one layer has passed, rearrangement of the layer may still be possible because the entire region has not yet fully developed. For example, 8 weeks after birth layer 4c in the visual cortex of the monkey is no longer affected by monocular deprivation, whereas the upper and lower layers continue to be susceptible for almost the entire first year...

(*Kandel, Schwartz, Jessell* 3d ed. 1991, p. 957)

105

## ABNORMAL COGNITION, NORMAL LANGUAGE

Williams Syndrome (deletion of several genes on Chr. 7, IQ of 50), is characterized by a distinctive, elfin facial appearance, along with a low nasal bridge and an unusually cheerful, sociable personality. Severe mental retardation, cannot tie shoe-laces, tell right from left, retrieve things from the cupboard, add two numbers, find their way, draw a bicycle. Understand complex sentences and fix ungrammaticalities. Fondness for unusual words.

Crystal (U. Bellugi): This is a story about chocolates. Once upon a time in Chocolate World there used to be a Chocolate Princess. She was such a yummy princess. She was on her chocolate throne and then some chocolate man came to see her. And the man bowed to her and he said these words ...

108

Model			
Wibom Age: 11;1 KBIT 70 (RA)			
Wibom Age: 1 KBIT 77 (AS)			
Control Age: 6;1 KBIT 122 (BD)			

109

## CONCLUSION

Language experience is a mere "trigger"

**Claim:** children only use linguistic experience to choose among a very narrow range of possible hypotheses given to them by the innate language module – universal grammar.

Plato's paradox is explained by the richness of the innate structures.

112

## INHERITED LANGUAGE DEFICIENCIES

Specific Language Impairment (KE family, 3 generations, 30 members, 16 affected, M. Gopnik)

Impairment uniform across affected members, randomly distributed; Impervious to teaching and correction; Persists through life; Problems with inflections (plural, tense, agreement); non verbal IQ tests normal, hearing normal.

- The boys eat four cookie
- Carol is cry in the church
- It's a flying finches, they are

110

## INHERITED LANGUAGE DEFICIENCIES

No plausible environmental causes

53% of family affected but only 3% of the total population

Syndrome all or none, so likely one gene is the cause rather than several genes, no graded disability

Autosomal gene – affects both sexes

Dominant gene (not recessive) – one copy

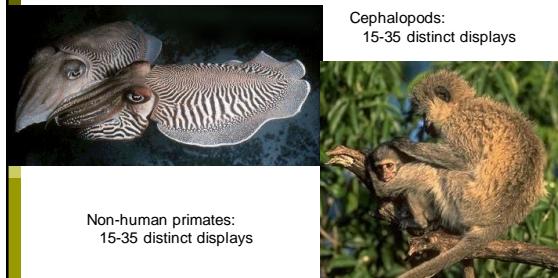
Led to the isolation of the FOXP2 gene by Lai et al which is implicated evolutionarily also in the leap to language

111

## Biology of Language

Where did human language come from, and why?

### After 450 million years...



## The Great Chain of Being

Vast chain of being! which from God began,  
Natures aethereal, human, angel, man,  
Beast, bird, fish, insect, what no eye can see,  
No glass can reach; from infinite to thee,  
From thee to nothing.

-Alexander Pope, *Essay on Man*

Humans are the pinnacle of development, highest in a chain of behavioural and communicative complexity

### Primates are “more evolved” than molluscs

- More complex bodies and brains
- More complex social structures
- More complex and flexible behavior
- Longer lived
- Better at learning and problem solving

BUT no real change in “vocabulary size”

In 1866, the Linguistic Society of Paris went so far as to ban debates on the subject (and also proposals for a universal language) — a prohibition which remained influential across much of the western world until late in the twentieth century.

### not just squid and monkeys...

For most relatively social adult fishes, birds and mammals, the range or repertoire size [of communicative displays] for different species varies from 15 to 35 displays.

-Encyclopedia Britannica, “Animal Communication”

- Little correlation between repertoire size and location in the "vast chain of being." Cuttlefish with as many different communicative displays as chimps.
- Selective (natural selection) pressures prevent the overall number of 'displays' from growing beyond a certain point, even though new displays are developed to suit new (evolutionary) adaptive circumstances.
- Similarly, although specialized physical organs develop in response to new evolutionary opportunities, other specializations tend to be lost, so that creatures do not over time accumulate indefinitely many humps, horns, claws, trunks etc.

We can ask: What aspects of the human genome make spoken language possible? What (natural) selective pressures on our ancestors led these characteristics to develop?

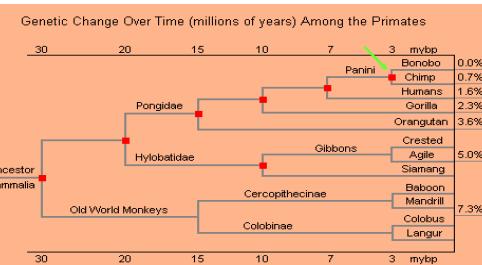
The human species has in fact adapted genetically to facilitate the use of spoken language. Thus it is worthwhile to look into what these adaptations are, and also at some theories about what selective advantage(s) they offered to our ancestors.

## Language and elephant trunks

- Human language, with its hundreds of thousands of words, is not just the **logical endpoint** of some obvious "scale of sensual, mental powers." Rather, it seems to be a behavioral counterpart of the peacock's tail or the elephant's tusk: a specific, enormously hypertrophied (overgrown) development of structures with different original functions.
- How and why did this happen? If complex systems of communication are so great, why hasn't evolution been developing them in other species for the last few hundred million years -- as eyes, ears, horns, claws etc. have repeatedly been developed?

## Monkeys-Man

From *The Third Chimpanzee* by Jared Diamond, 1992

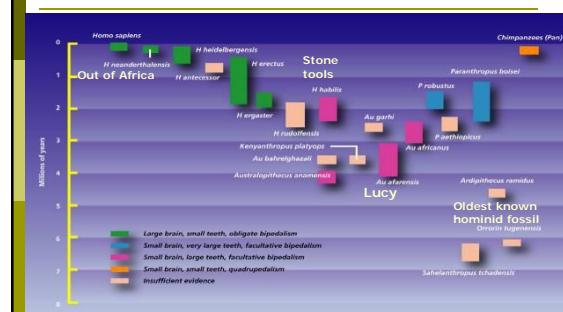


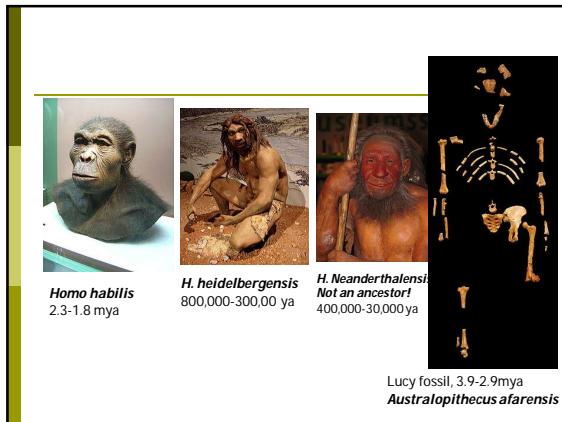
## Is language in our genes?

Is it meaningful to ask about the genetic evolution of language?

We share this *genetic language* with every other living thing on earth, but it is only to our fellow humans that we can talk! Although molecular genetics is not the kind of "language" under discussion, it provides one framework for interpreting our question about the origins of human spoken language.

## Hominids





## Vocal Tract Changes in Hominids

- The vocal tract changes occurred between the neck and the nose, and served to adapt our vocal tracts for speaking.
- Specifically, we shortened our muzzle and the oral cavity it contains, and stretched out our pharynx (throat) by lowering the larynx (the Adam's apple, voice-box).
- The result of these changes is to make it possible for our tongue to move forward and back, up and down, in a way that creates resonant cavities of different sizes in various places in the vocal tract.

## What hominids did

We are talking about evolution in the five million years since we separated from the ancestors of today's great apes (Chimpanzee, Gorilla, etc.).

The language-related changes took place from the neck up in two areas: the mouth and throat - the "vocal tract" - and the brain.

## Human and Chimpanzee Vocal-tract Anatomy

This image cannot currently be displayed

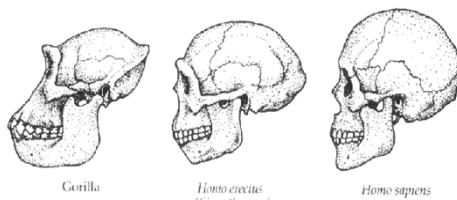
## We will leave aside several controversies

How many distinct species should be recognized in the fossil record (expert opinions vary from three to fifteen)?

- Where along the line from *Australopithecus* to *Homo erectus* (about 1.8 mya) to *Homo sapiens* (about 100,000 ya) did how much of the various changes take place?
- Where on the family tree do various particular species or subspecies (e.g. the Neanderthals) fit in? (mDNA suggests that they are not our ancestors, but did cross-breed)
- Did the recent change from *erectus* to *sapiens* happen in one place (the "out-of-Africa hypothesis") or over a wide area (the "multiregional hypothesis")?

## Skull comparisons

The skull of *Homo erectus*, our immediate ancestor (1.8 mya - 100,000 ya) appears to be intermediate in these respects

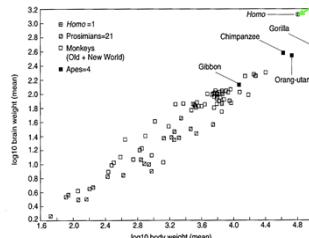


## Vocal Tract Changes

- These changes permit us to make a variety of speech sounds, but is otherwise a bad idea!
- Holloway 1996, *The evolution of the human vocal apparatus* says: The lower position of the larynx alters dramatically the way humans... breathe and swallow. The loss of the ability of the epiglottis to make contact with the soft palate means that the possibility of having two largely separate pathways, one for air and one for liquid, no longer exists. The respiratory and digestive tracts now cross each other in the area of the pharynx.. This new configuration can, and does, have unfortunate drawbacks. The major problem is that a bolus of food can become lodged in the entrance of the larynx. If this material cannot be expelled rapidly an individual may literally choke to death... Another disadvantage of the crossed pathways is the relative ease with which vomit can be aspirated into the trachea, and thus pass into the lungs.

## Brain changes

- Holloway 1996, *Evolution of the Human Brain* shows that the relationship of brain weight to body weight is roughly linear (on a log-log scale) across a large range of primate sizes. The data point for humans is above the trend line by a significant factor.

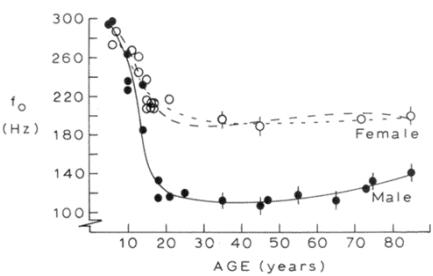


- This problem is even worse for men than for women, because as a secondary sexual characteristic of male humans, the larynx increases in size and moves lower in the throat at puberty. None of the other great apes show this laryngeal sexual dimorphism (differences between male and female), or indeed any other vocal tract dimorphism -though they have greater dimorphism in overall size, and also show dimorphism of canine teeth, which humans lack.
- The unique human development of sexual dimorphism in larynx size and position presumably means that vocalization is important to us in ways that it is not to gorillas and chimps.

## Structural brain changes

- There are four major reorganizational changes that have occurred during hominid brain evolution:
- (1) reduction of the relative volume of primary visual striate cortex area, with a concomitant relative increase in the volume of posterior parietal cortex, which in humans contains Wernicke's area
  - (2) reorganization of the frontal lobe, mainly involving the third inferior frontal convolution, which in humans contains Broca's area
  - (3) the development of strong cerebral asymmetries of a torsional pattern consistent with human right-handedness (left-occipital and right-frontal in conjunction); and
  - (4) refinements in cortical organization to a modern human pattern, most probably involving tertiary convolutions. (This last 'reorganization' is inferred; in fact, there is no direct palaeoneurological evidence for it.)

## Fundamental Frequency



The first three straightforwardly involve language in whole or in part. *Wernicke's area* in modern humans is involved in comprehension of language. *Broca's area* is involved in motor control of speech. The cerebral asymmetries in the third point involve a localization of language skills in the dominant (usually, left) hemisphere of the brain, and of other abilities (visuo-spatial and emotional) in the non-dominant hemisphere.

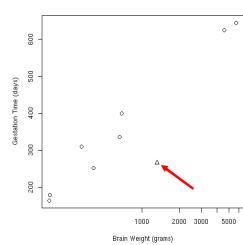
## Cost of Brain Changes

- ❑ Brain tissue is about ten times more expensive to maintain than other tissue. The human brain at 2% of our body weight, consumes about 20% of our energy.
- ❑ Increased brain size also translates to increased gestation period, because fetal brain tissue is laid down at a relatively constant rate. The graph on the next slide shows the relationship for a sample of species from rhesus monkeys to elephants.

## Signal systems

- ❑ **Chemical**, but will not travel far, nor upwind
  - ❑ **Neural signals** are not an option for communication between individuals (need to be plugged in).
  - ❑ **Tactile signals** require direct contact.
  - ❑ Signaling by manipulation of **electric fields** does not seem to be biologically reasonable, except for short-distance systems in water.
  - ❑ **Visual signals** are common and especially rich for many animals, especially primates but they are limited to line-of-sight and require visual attention on the part of the recipient.
  - ❑ The **acoustic channel** has an inherently lower capacity than the visual one does, but has the advantage of going around or through most obstacles, and being available regardless of the direction the organism is "looking".
- Hominid ancestors, when they split off from the lineage of chimps and gorillas (5mya) could have developed built-in sonar or an improved sense of smell. They might also have stayed about the same, as Homo Erectus, our immediate ancestor, did for two million years. Instead, they learned to talk.

## Brain and Gestation



- ❑ Triangle represents humans and if we bring it in line with the other species, human gestation should be about 15-17 months!
- ❑ Human infants are born premature and remain so for the first year of life.

## 3 Unique Features of Human Language

- ❑ Big, discrete vocabulary  
10,000-100,000 "words"... or more
- ❑ Recursive compositionality  
making bigger messages by combining smaller ones, more complex meanings by combining simpler ones
- ❑ Action to "change others' minds"  
we know others may have different knowledge and beliefs  
we communicate to inform, persuade, etc.

## Signaling universe

- ❑ Light, sound, air and water currents, chemicals, temperature gradients, all carry information about the structure of the world and the activities of its inhabitants.
- ❑ Creatures that can effectively access and use this information eat better, live longer, and reproduce more effectively, so there is *selective pressure* to develop sensors of all types.
- ❑ But pure sensors are useless: the organism has to evaluate the sensory information, act upon it and evaluate the feedback as well.

*Aside:* The basic mathematics of this integration of perception and action was worked out during WW II, motivated by the need for radar-guided anti-aircraft guns, auto-landing devices for airplanes, homing torpedoes etc. Fundamental work in this area was done by Norbert Wiener at MIT and by Andrey Nicolaevich Kolmogorov at Moscow State University (earlier by von Neumann). After the war, Wiener went on to develop the underlying metaphor of "control and communication" into the field of *cybernetics*. (Wiener's *The Human Use of Human Beings - Cybernetics and Society*)

## Many other "little" things...

- ❑ Displaced reference
- ❑ "Doubly digital" vocabulary (duality of patterning)  
words are discrete and well individuated  
words are patterns of digital sound elements ("phonemes")
- ❑ Variability in sound system and word meanings  
constant spontaneous social change -- new dialects  
adults have trouble adapting -- shibboleths
- ❑ Singing/chanting  
stylization of pitch and time in ratios of small integers
- ❑ Various specific formal properties
  - e.g. morphological "blocking"

## Principles of Evolution

- Consider the principles that underlie evolutionary reasoning in general:
- **Variation:** Variation in physiology/anatomy/behavior among individuals of a species
- **Heredity:** Offspring resemble their parents in traits more than they do unrelated individuals (typically for genetic reasons)
- **Natural Selection:** Individuals with certain traits leave more offspring than individuals with other traits, because possessing those traits makes them better able to do what they need to do to survive, reproduce, etc.

## Problems

- Problems
  - We have no 'close' relatives (e.g. 7-10 million years to a common ancestor with chimpanzees, vs. a few thousand for dogs and wolves)
  - We are not sure who our ancestors are
    - Evidence from the fossil record is limited
    - Inferences about cognitive abilities have to be based on e.g. cranial capacity (since stone tools are found in a number of hominid types)
    - There are no fossil records of behaviour

## Application to Cognition/Language

- If one is to provide an argument about how language evolved, several things have to be done
  - It must be shown that there is heritable variation in cognitive capacities
  - It must be shown that individuals with greater capacities leave more offspring
  - Etc.

## Look for Homologues

- Identical structures used for different functions: mammalian jawbones and human ear bones; mammalian fingers and bird wings etc.
- Brain areas have homologues between humans and monkeys/primates, and some functional connections can be studied (e.g. Broca's in macaques relates to facial movements)
- But it isn't clear what in a primate brain would be homologous to speech, since primates do not have speech
- Similar considerations apply in the attempts to teach language to non-humans; what we see could very well be analogy, in which case we wouldn't learn anything about the evolution of language

## Language is weird

- Quantitatively and qualitatively unique
  - like elephants' trunks
- No similar evolutionary trends in other species
  - other species don't "want" to pick up peanuts with their noses
    - all mammals have flexible noses, some use them as manipulators
    - no general trend to develop anything like trunks
  - other species don't "want" to exchange very complex messages
    - (nearly) all mammals make noises, some use them to communicate
    - no general trend to develop anything like human speech

3/26/2016

LING001

33

## Why talk? Speculation...

- The redesigned vocal tract is good for making different vocal sounds. The reorganized and expanded Broca's area deals with control of sound structures - *phonology* and *phonetics*. The reorganized and expanded Wernicke's area, along with the larger cortex in general, allows us to have lots of words, each one connecting a meaning with a pronunciation – *lexicon* and *semantics*. We also learned to think about what others believe – 'others minds' problem - which made us better at communicating regardless of the medium.

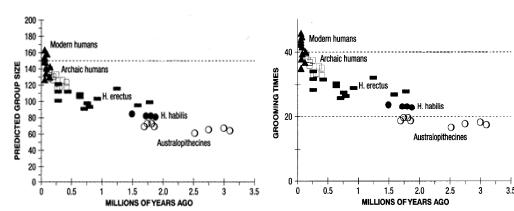
Terence Deacon, Harvard Biological Anthropologist writes: "...the myriad advantages that better communication could offer: organizing hunts, sharing food, communicating about distributed food sources, planning warfare and defense, passing on tool-making skills, sharing important past experiences, establishing social bonds between individuals, manipulating potential sexual competitors or mates, caring for and training young, and on and on." The linguist Derek Bickerton suggests that hominids invested in language so as to be able to think better. This hypothesis views rational thought as being at least in large part made up of inner speech. More recently, he proposes a striking variant of the "organizing hunts" theory: human language emerged because of the need to recruit and co-ordinate crews to help in scavenging the carcasses of naturally-deceased megafauna.

- Dunbar says that among primates, "encephalization" (brain size normalized for body size) varies in proportion to social group size. The larger the group a primate lives in, the more brain it needs to keep track of social relationships within the group. This is plausible, given the intricate micro-politics of primate society. If we assume that larger brains evolved in primates in order to permit larger social groups (e.g. for better intra-species competition or better defense against predators) we have what has been called the "Machiavellian Intelligence Hypothesis."
- If we look at human brain size from the perspective of this hypothesis, and extrapolate the relationship between brain size and social group size found in other primates, we predict a "natural" group size for humans of about 150.

## Any evidence?

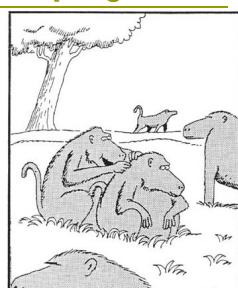
- The cited advantages certainly exist but are they strong enough or unique enough to explain why hominids developed language, while other animal lineages did not? In documented modern hunter-gatherer cultures, language does not play a very large role either in co-ordinating hunts or in teaching tool-making. Packs of wolves and wild dogs are extraordinarily good at group hunting, without being able to talk. Many kinds of human thought do not seem to involve language at all.
- The evolutionary process that got human language started must have been able to accomplish something special, even with a small, poor, stumbling kind of approximation to language that hominids would have been able to manage before any language-specific adaptations took place. That rudimentary language must confer enough of an advantage for natural selection to have a chance to create adaptations for complex, excellent language.
- Most recent theories assume that the crucial selective advantages of language were social. Perhaps something about the development of language made the creation and maintenance of larger social groups possible, at a time when larger social groups were essential to survival; or perhaps language permitted a different kind of social organization, enabling our ancestors to move into a different ecological niche.

## Group size and grooming time



## Grooming and Gossiping

- In his book *Grooming, Gossip, and the Evolution of Language*, Robin Dunbar proposes that our ancestors evolved language so as to use *gossip* as a more efficient substitute for the *grooming* behaviour that other primates use to establish and maintain social relationships.



- In primate societies, grooming (picking nits out of fur) is a major factor in establishing and maintaining social bonds. The bigger the primate group, the more time on average each member spends in grooming others. If we look at human social relations in this perspective, with a group size of 150, we should have to spend 40% of the day in grooming. This is far too high to be practical - the highest actual proportion observed among primates is 20% (in certain baboons).
- Dunbar suggests that our ancestors, facing hard times on the African plains, very badly needed to live in larger groups. "Gossiping" (in whatever form it first arose) made it possible to form and maintain social bonds more efficiently than grooming, both because more than two can do it at once, and also because you can do some useful work (gathering or processing food) at the same time. The development of sense and reference - and especially of proper names for group members - enabled political maneuvering at a higher level in larger groups.

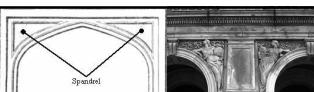
## Marriage Contracts

- In *The Symbolic Species*, Terrence Deacon argues that hominid brains and human language have co-evolved over the past two million years, driven by "a reproductive problem that only symbols could solve: the imperative of representing a social contract", which in turn was required to take efficient advantage of the resources available via systematic hunting and scavenging for meat.
- Deacon thinks that early hominids developed symbolic communication as a way to establish social contracts permitting stable family and group structures, which otherwise would not have permitted hunting and scavenging for meat as a systematic source of supplemental food during times of drought. This set the state for nearly two million years of evolutionary adaption for improved symbolic communication, probably due to sexual selection (crudely, females preferred males who could make more convincing promises).

## Conclusions

- Unlike bones, behaviour leaves little evidence in the fossil record. Since no other species has developed a symbolic communication system like human language, we are not in a good position to make generalizations, except about the many cases where symbolic language did not develop. Therefore, it is difficult to make a strong case for or against the various theories of the evolutionary precursors and selective advantages of human language.

## Spandrels



- Another perspective on the initial development of language treats it as a sort of accidental side-effect of larger brains, which on this view developed for some other reason (say to facilitate tool use and/or social dynamics).
- This "side-effect" theory would be an example of what Stephen Jay Gould has called evolutionary *spandrels*. The original meaning of "spandrel" is a space between two arches and a horizontal cornice above them; this space began as an accidental (but unavoidable) consequence of architectural techniques based on the use of arches and domes; because this accidental space is a convenient place to put paintings, it developed into a planned part of buildings with a specific function. Gould argues that many evolutionary developments are of this kind - some feature arises as an accidental side-effect of another change, but then turns out to be useful and comes to be itself shaped by selective pressures.
- This spandrel theory is not inconsistent with other accounts of the selective pressures for language development.

## Next

- What do we know about other animal communication systems? Can we 'teach' language to other animals?

## If language is so great, why doesn't every species get one?

Possible answers:

- It's too expensive, relative to the benefits
  - e.g. in terms of brain tissue requirements
- It's hard to get started
  - e.g. requires an unlikely evolutionary "invention" – it is not just an extension of animal communication systems
  - or, early "releases" are not very useful since "theory of mind" may be lacking or displaced reference can be confusing

## Animal Communication

Same as or different from human languages?

### Where to look?

Social animals  
Large brains (relative to body size)  
Our close relations  
Visible or audible modalities  
Accessible environments  
Not too aggressive!

### Human language

innate	critical period
arbitrary	compositional
structured	recursive
infinite	displaced reference
flexible	creative

Systems with ALL these properties are termed language in the narrow sense

### Species that qualify on one or more grounds

Honeybee  
Various birds  
Some monkeys  
Great apes  
•Cetaceans

### Question

Do animal communication systems show these properties?

1. naturally intra-species in the wild?
2. when taught inter-species in the lab?

### Design features

1. Use of Vocal-Auditory Channel
2. Arbitrariness
3. Semanticity
4. Cultural Transmission
5. Spontaneous Usage
6. Turn-taking
7. Duality
8. Displacement
9. Structure Dependence
10. Creativity (openness)

## Animal Communication

- ❑ Animal communication systems have a limited *signal set*: the inventory of things that are communicated about is very limited (e.g. food, sex, aggression, predators/threats)
- ❑ Human language, as you may recall, has the ability to incorporate new basic elements (typically words) in a way that is unlike what is found in animals
- ❑ Moreover there is nothing like *syntax* for generating new 'utterances' in the animal kingdom (in some domains linear order is important, however, e.g. birdsong)

## Displacement

- ❑ What the reading refers to as *situational freedom* or *displacement* involves referring to things that are not immediately present
- ❑ This is an obvious property of human language; but with limited exceptions, this is clearly not the norm in animal communication

- ❑ Consider a simple example of communication: dog posture. In this case, the dog is communicating that it is ready to play.



## Some examples: Bees and Vervets

- ❑ **Vervet Monkeys:** different calls for different predators, which trigger appropriate behaviour in the listener
- ❑ **Bees:** Communicate distance, richness and location (relative to the hive) of food sources (using the sun as a fixed point of reference) through dance

- ❑ The dog in the example above is communicating something, namely, something about its internal state
- ❑ With *reference*, we mean roughly what it is that a signal in communication refers to
- ❑ Most animal communication systems are limited in what their signals are capable of referring to; typically, an internal state, or an immediately present stimulus

## Vervet Monkeys

- ❑ **Snake Alarm:** Hearers stand on hind legs and look on the ground
- ❑ **Leopard Alarm:** Hearers run to the top of the nearest tree (where leopards can't go)
- ❑ **Eagle Alarm:** Vervets run under a bush/cover of tree branches

## Function

- The organization of the alarm system makes sense given the living conditions of the vervets
- The predators are categorized in different ways: especially, in ways that require different responses
- Categorization is acquired in the first four years of life: young vervets miscategorize, but eventually learn the correct association of calls with predators (without, apparently, any correction)

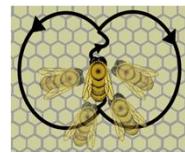
## Other dances

- The Italian honeybee also has:
  - Round dance = near hive (within 20 feet or so)  
Gives information about distance, and quality (more vivacity)
  - sickle dance = intermediate distance, 20-60 feet

- Note that there are several ways in which to interpret what the different calls might 'mean', if we approach it that way
  - 'Leopard!'
  - 'Head for the trees!'
  - Don't stay on the lower branches
- Even in systems of this type, which show some complexity because of the categorizations involved, the calls are made in the *presence* of the stimulus

## The Waggle Dance

- The dance proceeds in a figure 8 pattern
- The orientation of the dance with respect to either vertical or the position of a light indicates the direction (other factors indicate approx. distance)

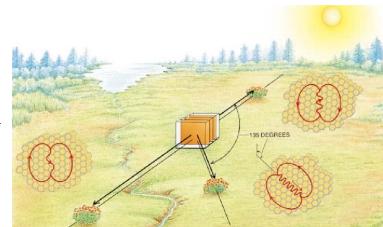


## Bee Dancing

- Honeybees forage for food sources and, upon returning from a successful trip (i.e. after finding a food source) they perform a dance
- This dance is called the *waggle dance* because it involves this particular motion
  - The waggle dance conveys aspects of the journey the bee has completed- in particular, the location and distance of the food source
  - Other bees that have witnessed the dance then go to the food source

## Orientation

- Dances oriented directly to vertical indicate that food is in line with the sun
- Otherwise, orientation of the dance indicates the angle of the food from the sun



- Is the dance really signaling food? Or it other information like chemical signals that indicate the food source? This can be tested by making bees 'lie'.
- A further fact about the bee dance
  - When it is dark in the hive, the dance is oriented towards the vertical
  - When there is a light source visible in the hive, the dance is oriented towards the light source

## Birds

- Most bird calls innate, but not all
- Researchers in the 1950s raised a bullfinch with a canary.
  - Bullfinch learned the canary's song.
  - Its offspring (& its offsprings' offspring) did so too



- Bees detect overall light with their *ocelli*, a set of photoreceptors on top of the head
- Experimenters can paint the ocelli of bees, so that bees with painted ocelli behave as if it is dark, whereas bees with unpainted ocelli act as if it is light
- The experiment involves bees with painted ocelli dancing about the location of a food source; the dance is witnessed by bees with unpainted ocelli

- Chaffinch have dialects
  - As 1-year-olds, they learn the dialect of the area in which they have chosen to live and breed.



- **Dancing Bees:** The dancing bees orient with respect to vertical
  - **Witnessing Bees:** The witnessing bees interpret the dance with respect to a light source, which is not at vertical
  - Result:** Witnessing bees (the majority) went to the source that was indicated by the dance, not the source actually visited by the dancing bees (who were giving false information)
- This is consistent with the communication hypothesis, but not the odour-alone hypothesis.**

- A chaffinch reared in isolation had a simple, proto-song. The right length, but without the typical structure.
- Like human language?



## Human language in non-humans

- Another aspect of research on language in animals involves a different methodology than studying communications systems that are natural to non-human species
- This involves trying to teach non-human animals (typically chimpanzees, bonobos, or e.g. gorillas in the case of Koko) human language

## Basic Results

- In the basic case, it seems clear that the primates are able to associate signs (whether visual or manual) with a 'meaning'
- This is unsurprising given that many species can associate a symbol of some type with something else (consider e.g. dogs)
- Whether this amounts to the primates having a 'vocabulary' is a more difficult question

## On the methodology

- To some extent, studies of this type capture the popular imagination, often concentrating on the question of whether animals have 'true language'
- Snowdon (1993) has a perspective on this:  
"...ethologists who study natural communication differ from psychologists who try to teach human language to other species in the same way as anthropologists differ from missionaries.  
Anthropologists try to understand the natives whereas missionaries try to civilize them."

## Harder questions

- An additional question is whether the primates who have been studied are able to use new symbols in new combinations (creativity, rule-governance)
- This is, of course, a central facet of human language
- In many cases, claims for new use of symbols or 'proto syntax' are taken from long sessions, and are not always representative
- For instance, it has been claimed that Lana, upon encountering an orange for the first time, signed: "apple which-is orange-color"

## Some facets of the programme

- In most of these cases, the primates were taught some version of sign language, such as with Washoe
- In a few cases, the primates were taught visual symbols
- There are some differences as well in the training situation (home, laboratory, etc.)



## More context

- Lana
- Tim: What color of this ?  
Lana: Color of this orange
  - Tim: Yes
  - Lana: Tim give cup which-is red
  - Tim: Yes
  - Lana: Tim give which-is shut ?  
Shelley give ?
  - Tim: No Shelley
  - Lana: Eye // Tim give which-is orange?
  - Tim: What which-is orange
  - Lana: Tim give apple which-is green ?
  - Tim: No apple which-is green
  - Lana: Tim give apple which-is orange ?
  - Tim: Yes
- (Rumbaugh)



Note the use of fixed expressions and the fact that the trainer introduced color into the conversation



## Synopsis

- The idea that human language should be taught to non-humans involves several components
  - One the one hand, it is asking primates to do something that they do not by nature do
  - This seems to relate to a perspective on human language which we have found evidence against before: if language is just general intelligence, then primates and other 'intelligent' species should be able to learn some of it
  - Of course, we might learn something about primate intelligence in such projects, but what we learn about human language is limited

## Conclusions

- In complex systems of animal communication, there are questions about
  - What animal calls, etc. 'refer to' (if they refer)
  - Whether or not they are used to influence the behavior of others, or are simply responses to stimuli
  - Some evidence exists for the encoding of information about stimuli that are not present (bees)
- It is unclear what can be learned about language per se from experiments that try to teach human language to non-humans

### **Summary: Trees**

- (i) In every sentence, certain groups of words form 'natural units' [=constituents] and may:
  - stand alone
  - be moved as a unit
  - be replaced as unit by a pronoun
- (ii) Trees encode the information about constituents: two expressions are a natural unit (=constituent) if there is a sub-tree that contains them and nothing else.
- (iii) A sentence that can be analyzed as 2 different trees is structurally ambiguous (e.g. *Lucy will hit the student with the book*)

### **A Puzzle Explained: Question Formation**

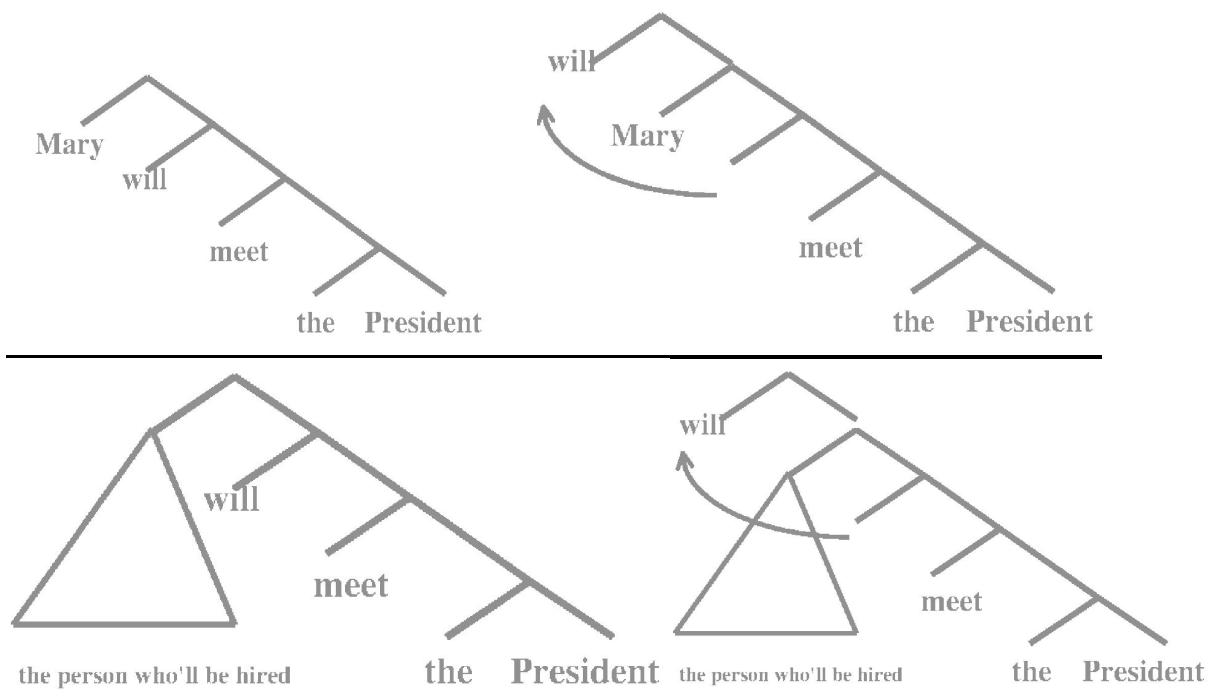
If we wish to form a question that corresponds to the assertion *John is in the garden*, we may simply move the auxiliary *is* to the beginning of the sentence, yielding *Is John \_\_ in the garden?* In a slightly more complex case, as in (1a), we form the corresponding question by moving to the beginning of the sentence the first *is*, yielding (1b). If we tried instead to move the second *is*, we would obtain a sharply ungrammatical result as in (1c):

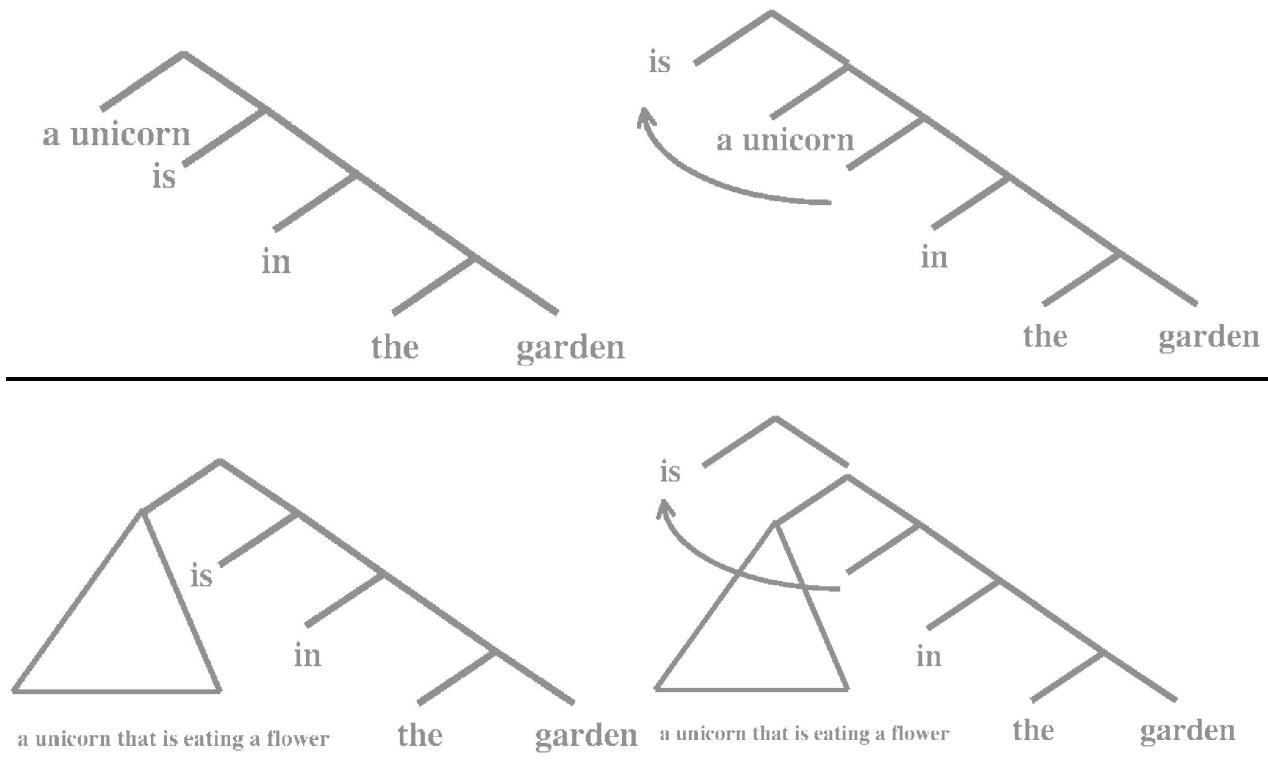
- (1) a. John is in the garden next to someone who is asleep.  
 b. Is John \_\_ in the garden next to someone who is asleep? (Move the first *is*)  
 c. \*Is John is in the garden right next to someone who \_\_ asleep? (Move the second *is*)

The rule of question formation cannot be to move the auxiliary that is first/second in line, see 2:

- (2) a. A unicorn that is eating a flower is in the garden  
 b. \*Is a unicorn that \_\_ eating a flower is in the garden? (Move the first *is*)  
 c. Is a unicorn that is eating a flower \_\_ in the garden? (Move the second *is*)

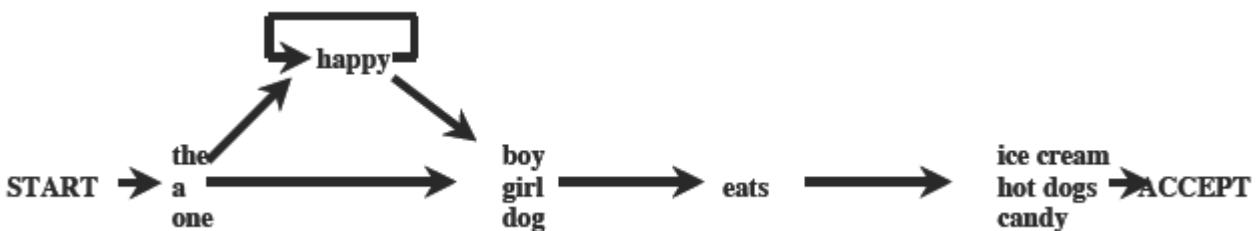
As we have seen the rule is structurally defined as movement of aux to a stipulated position COMP across a certain constituent (subject). A child that only heard simple cases of question formation (e.g. *Is John \_\_ in the garden?*) would have to infer a rather complex and subtle rule from limited data.





### An Incorrect Model: Finite State Grammars (=Markov Model)

A plausible -but incorrect- model is (also discussed by Pinker in Chapter 4 of *The Language Instinct*, available in Central Library) the Finite State Model (also called 'Markov Model'; Pinker also calls it a 'word chain device'). It is both natural and historically important, since it was considered plausible until the 1950's. In a nutshell, it attributes to a speaker a simple mental system that allows him or her to determine whether a given word can or cannot follow another given word. Here is the example of a Finite State Model discussed by Pinker (with 'START' and 'ACCEPT' states, which are implicit in Pinker's discussion); the idea is that you feed the sentence to the machine, starting with the first word, one word after the other. If you end up in the ACCEPT state after the last word has been processed, the sentence is accepted; otherwise the sentence is rejected):



Examples of sentences that are generated by above:

- the boy eats ice cream
- the happy boy eats ice cream
- the happy happy boy eats hot dogs
- a happy happy girl eats candy

Examples of ungrammatical sentences that are not generated by above:

- \*boy the eats ice cream
- \*happy boy eats hot dogs
- \*hot dogs eats the dog

Examples of grammatical sentences that are not generated by above:

- a. some boy eats ice cream
- b. the dog that the dog eats eats ice cream
- c. either the boy eats ice cream or the girl eats candy

There are two important arguments against the Finite State Model:

**Argument 1:** It does not account for the tree-like structure of sentences that we have observed-

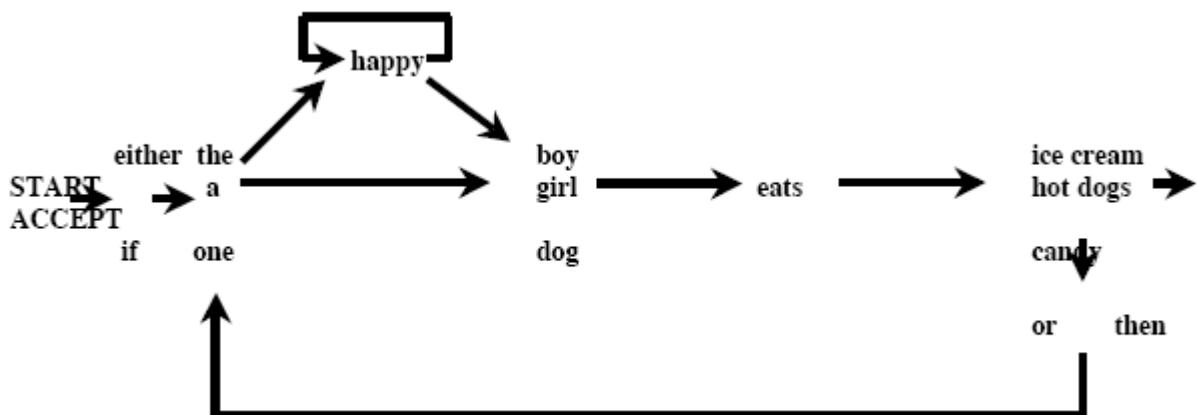
**Argument 2:** It cannot properly account for 'long distance dependencies', i.e. constructions in which two elements that depend on each other are separated by an arbitrary number of words.

Example of a long distance dependency: either ... or...

- a. Either John is sick or he is depressed
- b. Either John thinks that he is sick or he is depressed
- c. Either Mary knows that Johns thinks that he is sick or she is depressed
- d. Either the boy eats hot dog or the dog eats hot dog
- e. Either the happy happy boy eats hot dog or the dog eats candy

and so on...

We could try to integrate the *either ... or* construction into our Finite State Model, but no simple solution would work. To see this, observe that in the following model nothing requires that a sentence that starts with *either* should also contain *or* somewhere down the road. And for good reason: in order to 'remember' this, the model would need some kind of memory, which it lacks completely. The problem turns out to be very severe. In fact, Noam Chomsky became famous in the 1950's by *proving* that no matter how complex a finite state machine was, it could not handle all constructions of English and thus that Natural Languages are a different kind of system altogether.




---

Some grammatical sentences generated by the above

- a. Either a girl eats candy or a boy eats hot dogs
- b. Either a happy girl eats candy or a boy eats hot dogs

Some ungrammatical sentences generated by the above

- a. \*Either a girl eats candy
- b. \*Either a happy girl eats candy

### Phrase Structure Grammars meet the requirements:

**Requirement 1:** Account for the tree-like structure that sentences have, and

**Requirement 2:** Provide an analysis of long-distance dependencies, i.e. constructions in which two elements that depend on each other are separated by an arbitrary number of words.

**Example of a UG parameter:**

Consider the data below from French, English, Italian and Catalan (\*indicates ungrammaticality)

- |                 |           |
|-----------------|-----------|
| (1) a. __ parla | (Italian) |
| b. __ parla     | (Catalan) |
| c. *__ parle    | (French)  |
| d. *__ speaks   | (English) |

Property 1: Italian and Catalan allow verbs to have a *null subject*; French and English don't.

Property 2: In the table below, are the person endings (on verbs) in these languages (the French pronunciation is given, not the spelling). Italian and Catalan have different forms across various persons but not English and French:

	1 <sup>st</sup> singular	2 <sup>nd</sup> singular	3 <sup>rd</sup> singular	1 <sup>st</sup> plural	2 <sup>nd</sup> plural	3 <sup>rd</sup> plural
<b>Italian</b>	parl-o	parl-i	parl-a	parl-iamo	parl-ate	parl-an(o)
<b>Catalan</b>	parl-o	parl-es	parl-a	parl-em	parl-eu	parl-en
<b>French</b>	parl	parl	parl	parl-ō	parl-é	parl
<b>English</b>	<b>speak</b>	<b>speak</b>	<b>speaks</b>	<b>speak</b>	<b>speak</b>	<b>speak</b>

Property 3: Subjects precede the verb in all 4 languages but may follow the verb in Italian & Catalan but not in French & English:

- |                                |           |
|--------------------------------|-----------|
| (2) a. __ ha telefonato Gianni | (Italian) |
| b. __ ha telefonat en Joan     | (Catalan) |
| c. *__ a téléphoné Jean        | (French)  |
| d. *__ telephoned John         | (English) |

Property 4: Embedded questions with the question word in the subject position are permitted in Italian and Catalan but not in French and English:

- |  |           |
|--|-----------|
| (3) a. Chi credi che telefonerà?           | (Italian) |
| b. Qui creus que telefonarà?               | (Catalan) |
| c. *Qui crois-tu que téléphonera?          | (French)  |
| d. *Who do you think that will telephone ? | (English) |

The English question has the question word *who* which is the subject of the verb *telephone*:

You think that *who* will telephone

This question is permissible when you are seeking clarification (having not heard or misheard), but the question form as in 3d is not acceptable to native speakers of English. If the question word originated in the object position, there would be no such problem:

- |  |           |
|--|-----------|
| (4) a. Qui crois-tu que Michelle rencontrera __? | (French)  |
| <i>Who believe you that Michelle will-meet ?</i> |           |
| b. Who do you think that Michelle will meet __?  | (English) |

- (5) a. \_\_\_\_\_ you think that Michelle will meet **who**?



- b. **who** (do) you think that Michelle will meet \_\_ ?



- (6) a. \* \_\_\_\_\_ you think that **who** will telephone?



- b. \***who** (do) you think that \_\_ will telephone ?



The (interesting) point is that the question word when it originates from the position before the verb is allowed in Italian and Catalan but not in French and English.

The properties above are summarized below:

	<b>Null Subject</b>	<b>Person Morphology</b>	<b>Subj after V</b>	<b>Preverbal questions</b>
<b>Italian</b>	Yes	Yes	Yes	Yes
<b>Catalan</b>	Yes	Yes	Yes	Yes
<b>French</b>	No	No	No	No
<b>English</b>	No	No	No	No

The Null Subject parameter was postulated to account for this cluster of properties. A single (abstract) parameter accounts for a range of differences (language variation) but constrains the variation quite severely. (Yes to all properties entails a null subject language). The theory predicts that there ought to be no ‘mixed’ cases. If this is true, for children to induce that their language is a NS language or not, they need evidence from any one property.

Why this set of four properties? Linguists seek explanation for the relation between these properties. For example, 2 may entail 1; and 4 may be a consequence of 3 etc.

## Language Change: Diachronic or Historical Linguistics

### Languages Change

All the time. How and why?

1. Language learning
2. Language contact
3. Change for social reasons
4. Natural linguistic processes

Is this good or bad? Who cares?

- All stages of language are valid expressions of our language instinct (Universal Grammar)
- Just as all languages and dialects are valid expressions of our language instinct

### Sir William Jones

(28 September 1746 – 27 April 1794, Asiatic Society Founder)



□ The *Sanskrit* language, whatever be its antiquity, is of a wonderful structure; more perfect than the *Greek*, more copious than the *Latin*, and more exquisitely refined than either, yet bearing to both of them a stronger affinity, both in the roots of verbs and the forms of grammar, than could possibly have been produced by accident; so strong indeed, that no philologer could examine them all three, without believing them to have sprung from some common source, which, perhaps, no longer exists; there is a similar reason, though not quite so forcible, for supposing that both the *Gothic* and the *Celtic*, though blended with a very different idiom, had the same origin with the *Sanskrit*; and the old *Persian* might be added to the same family.

(Address to the Asiatic Society of Calcutta 1786)

### Language Learning

- Remember the typical processes of language acquisition: (Over-) regularisation (*scant* was *scamt*) Simplification
  1. The child must construct their language based on the input received
  2. This process is imperfect
  3. Bias towards regularization – learning an irregular form requires more input
  4. Also random differences may spread, especially through a small population

### Pateroster/Lord's prayer

- Latin liturgical version:  
Pater noster, qui es in caelis:  
sanctificetur Nomen Tuum;  
fiat voluntas Tua,  
sicut in caelo, et in terra.  
FADER USÆR ðu arðin heofnu  
Sie gehalgad NOMA DIN.  
Tocymeð RÍC DIN.  
suæ is in heofne and in eorðo.  
□ Northumbrian (c.650AD)  
Our fadir that art in heuenes, halewid be  
thi name; thi kyndoom come to: be thi wille  
don in erthe as in heuene:  
□ Old English (c.1100) *Matt. 6.9*  
Fæder ure þu be eart on heofonum; Si þin  
nama gehal god to become þin rice gewurþe  
ðin willa on eorðan swa swa on heofonum.  
□ Middle English (c. 1380)  
Oure fadir that art in heuenes, halewid be  
thi name; thi kyndoom come to: be thi wille  
don in erthe as in heuene:  
□ Early Modern English (c. 1559), *Book of Common Prayer*  
Our Father which art in heaven, hallowed be  
thy name. Thy kingdom come. Thy will  
be done in earth as it is in heaven.  
□ Late Modern English (BCP 1928)  
Our Father, who art in heaven, hallowed be  
thy Name. Thy kingdom come. Thy will be  
done, On earth as it is in heaven.

### Language Contact

- Entirely new languages can be created this way (Pidgins and Creoles)
- More commonly, new 'bits' are added to existing languages which don't change their identity (immediately)
  1. Through migration, conquest, trade
  2. Adults may learn the new language as a second language
  3. Children may be fully bilingual
  4. Results in borrowing of words, sounds, even syntactic constructions

## Tok Pisin (Papua New Guinea)

### The Lord's Prayer

Papa bilong mipela Yu stap long heven.

Nem bilong yu i mas i stap holi.

Kingdom bilong yu i mas i kam.

Strongim mipela long bihainim laik bilong yu long graun,  
olsem ol i bihainim long heven.

	Singular	Dual	Trial	Plural
1st exclusive	mi (I)	mitupela (he/she and I)	mitripela (both of them, and I)	mipela (all of them, and I)
1st inclusive	-	yumitupela (thou and I)	yumtripela (both of you, and I)	yumpela or yumi (all of you, and I)
2nd	yu (thou)	yutupela (you two)	ytripela (you three)	yupela (you four or more)
3rd	em (he/she)	tupela (they two)	tripela (they three)	ol (they four or more)

## Natural linguistic processes

### Some examples:

- ❑ Assimilation: /NPR / → [m pi ar] → [m pi ar]
- ❑ Dissimilation: /haws/ + [PL] → [haws★z]
- ❑ Syncope (deletion): /lajbr★ri/ → [lajb★ri] → /tɔi:z stejk/ → [tɔi:stejk]
- ❑ rapid or casual speech produces assimilation, vowel reduction, deletion
- ❑ this pronunciation can become conventionalized, and so end up being produced even in slower, more careful speech

## Borrowing

Arabic:	admiral, assassin, mask, mattress, racket, syrup, zenith
Chinese:	junk, ketchup
Czech:	robot
Gaelic:	dad, slogan, trousers
Greek:	biology, coma, method, science (and many other scientific and technical terms)
Hebrew:	camel, ethnic, jubilee, paradise
Hindi:	bungalow, pundit, shampoo
Italian:	balcony, ballot, jeans, traffic, volt
Native American:	caucus, cockroach, squash
Persian:	arsenic, lilac, pajamas, sherbet, talisman
Portuguese:	buffalo, caste, port
Russian:	bistro, disinformation, mammoth, sputnik
Sanskrit:	candy, jungle, swastika
Spanish:	banana, cannibal, cork, potato, sherry
Swedish:	gauntlet, slug, weld
Turkish:	coffee, kiosk, scarlet, sorbet, yogurt
Yiddish:	bagel, chutzpah, kosher, schmooze, schmuck

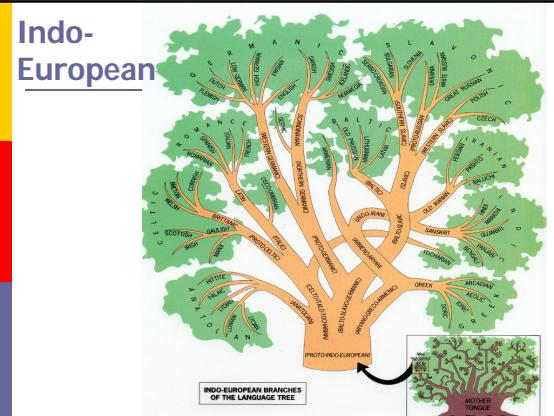
## Classifying Languages

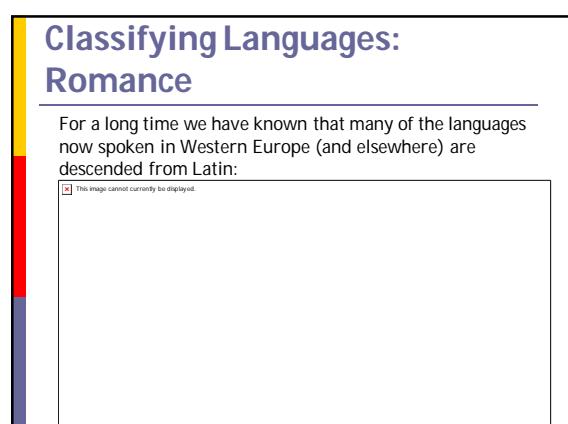
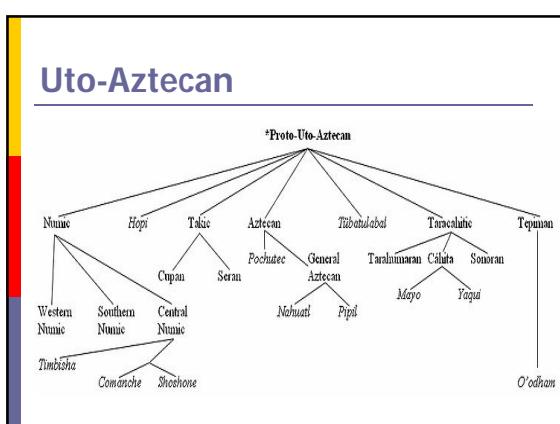
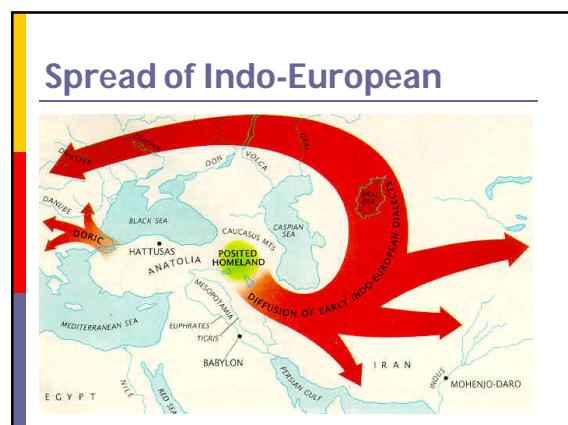
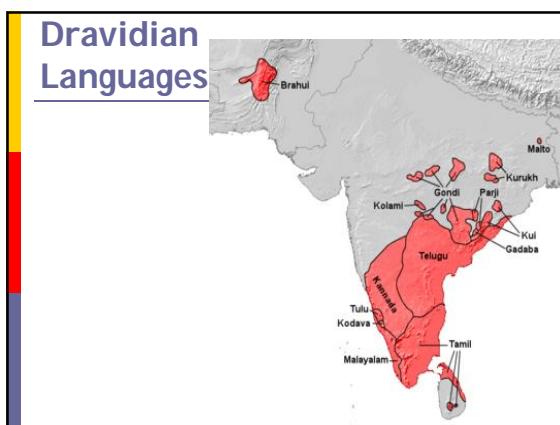
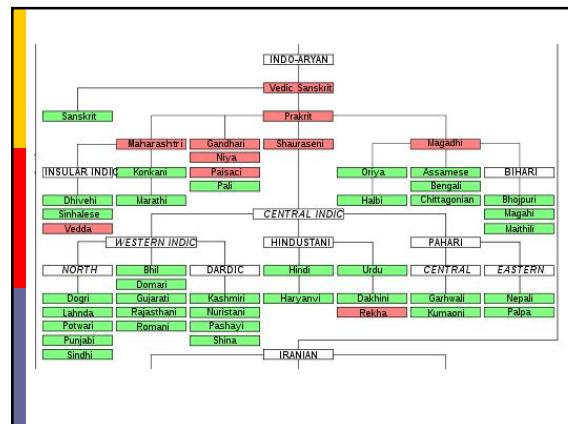
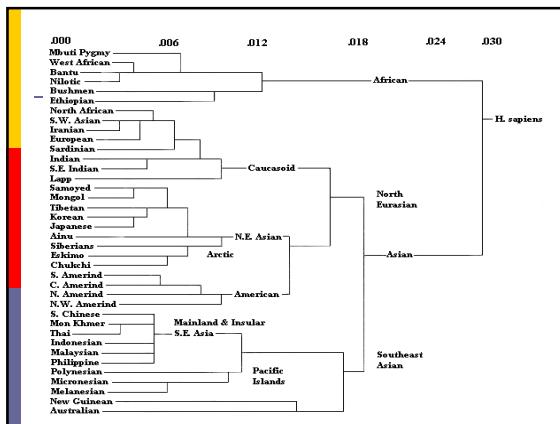
- ❑ Either because of what we know or because of what we can work out, we can classify languages according to their *origins*.
- ❑ Kinship terminology is used and languages are put into **families** (and sub-families) the relationships between languages are described using female familial terms: most often **daughter** and **mother**

## Change for social reasons

- ❑ People change the way they speak according to their social status or the way they would like it to be seen
- ❑ Educational policies

## Indo-European





## Classifying Languages: Germanic

2000-500	500-1 BC	1-500 AD	500-1000	1000-1500	1500-2000
Proto-Germanic East		Gothic		Crimean Gothic	
		Vandalic			
Runic	North	Old Norse	Old Icelandic Old Norwegian Old Swedish Old Danish	Icelandic Norwegian Swedish Danish	
	West	Old High German	Middle High German	German	
				Swiss German Pennsylvania Dutch Yiddish	
		Old Saxon	Middle Low German	Low German	
		Old English	Middle English	English	
		Old Dutch	Middle Dutch	Dutch	Afrikaans

## Cognates, more distant

	Sanskrit	Cl. Greek	Latin
1	éka	hei:s	u:nus
2	dvá	dúo:	duo
3	trí	trei:s	tre:s
4	catúr	téttares	quattuor
5	pánca	pénte	quinque
6	ṣaṣ	héx	sex
7	saptá	heptá	septem
8	aṣṭá	októ	octo:
9	návra	ennéa	novem
10	dáṣa	déka	decem

## Paternoster revisited

### Afrikaans

Ons Vader wat in die hemel is,  
laat u Naam geheilig word;  
laat u koninkryk kom;  
laat u wil ook op die aarde geskied,  
net soos in die he

### Danish

Vor Fader, du som er i Himmelne!  
Helliget vorde dit navn;  
komme dit rige:  
ske din vilje  
på jorden, som den sker i Himmelnen;  
mel.

### Dutch-Flemish

Onze Vader in de hemel,  
uw naam worde geheiligd,  
uw koninkrijk kome,  
uw wil geschiede,  
op aarde zoals in de hemel.

### German

Vater unser im Himmel,  
Geheiligt werde dein Name.  
Dein Reich komme.  
Dein Wille geschehe, wie im Himmel so auf Erden.

### Icelandic

Fa-ir vor, Þú sem er á himnum.  
Helgist Þitt nafn, til komi Þitt riki,  
ver-i Blinn vilji, svo á jor-u sem á himni.

### Swedish

Vår fader, du som är i himlen.  
Låt ditt namn bli helgat.  
Låt ditt rike komma.  
Låt din vilja ske,  
på jorden så som i himlen.

## Language Classification: How?

We rely on two things:

- the **Uniformitarian Principle**
- the **absolute regularity of sound-change**

## Cognates

English	Dutch	German	Danish
one	een	ein	en
two	twee	zwei	to
three	drie	drei	tre
four	vier	vier	fire
five	vijf	fünf	fem
six	zes	sechs	seks
seven	zeven	sieben	syv
eight	acht	acht	otte
nine	negen	neun	ni
ten	tien	zehn	ti

## The Uniformitarian Principle

'knowledge of processes that operated in the past can be inferred by observing ongoing processes in the present'

or, for language:

'Language must work now in the same way as it ever did'

## Regularity of Sound-Change

Most of historical linguistics relies on the assumption that

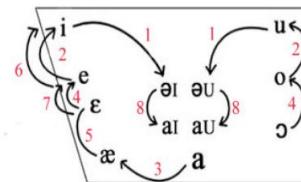
**sound-change is regular and without exception**

That is, any sound-change will affect **all** the words that contain that (combination of) sound(s).

## Regularity of Sound-Change

The **Great Vowel Shift** (primarily between 15-16<sup>th</sup> c.); the big difference between Chaucer and Shakespeare

the long stressed vowels of English underwent a major shift:



## English Orthography

[i] beet, beat	through, boot [u]
[ɪ] bit	put [ʊ]
[ey] bait, mate	dough, low, boat [ow]
[e] bet	[ʌ] but
[æ] bat	fall, caught [ɔ]
bite [ay]	[ɑ] cot
	[aw] about

□ 5<sup>th</sup> century Germanic invasions replaced Latin and the Celtic languages with Germanic Dialects (Anglo-Saxon: Angles, Saxons and Jutes)

□ Changes in Anglo-Saxon are usually grouped as:

**Old English (450-1100)**

**Middle English (1100-1500)**

**Early Modern English (1500-1800)**

**Present-Day English (1800-present)**

## Shouldn't it be...?

i bite	but u
ey beet	boot ow
a bat	
bait ay	aw caught

## Chaucer vs. Shakespeare

And whan I swoghe he wolle never fine  
To reden on this cursed boke al night,  
Al soðenly three leves have I plight  
Out of his boke right as he redde, and eke  
I with my fist so took him on the cheeke  
That in ourse fir he fli backwardes adown.  
And up he sterte as dooth a wood leon  
And with his fist he smoot me on the heed  
That in the floor I lay as I were dead.  
And whan he swaghe how stille that I lay,  
He was agast, and wolde have fled his way  
Till atte faste out of my swoghe I braide:  
"O hastou slain me, false thief!" I saide,  
"And for my land thus hastou mordred me?  
Er I he deid vit wol I kisse thee."

Enter Hamlet.  
Cer. Madane, will it please your grace  
To laue vs here?  
Qu. With all my hart. exit.  
Cer. And here Ofelia, ready on this booke,  
And walke aloofe, the King shal be vnfene.  
Ham. To be, or not to be, I here's the point,  
To die, to sleep, is that all? I all.  
No, to sleepe, to dreme, I may there it goes,  
For in that dreme of death, when wee're awake,  
And borne before an everlasting Judge,  
From whence no passenger ever returnyd,  
The vndiscovered country, at whose sight  
The happy smile, and the accyfled damn'd.  
But for this, the ioyfull hope of this,  
Who'd bear the scornes and flattery of the world,  
Scorned by the right rich, the rich curst of the poore

## Long vowels of Middle English

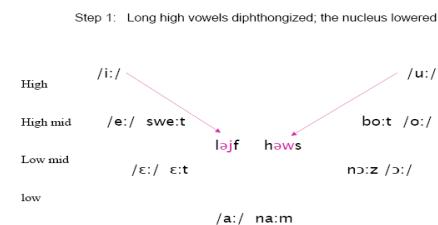
High	/i:/ lif	hu:s /u:/
High mid	/e:/ swē:t	bo:t /o:/
Low mid	/ɛ:/ ē:t	nɔ:z /ɔ:/
low	/a:/ nā:m	

It's only the long stressed vowels that were involved in the Great Vowel Shift, so we'll only show them in the following slides.

## High-mid vowels raise

High	swit	but
High mid	/e:/	/o:/
Low mid	ləf həws	nɔ:z /ɔ:/
low	/ɛ:/ ē:t	/a:/ nā:m

## Stepwise shift



## The result

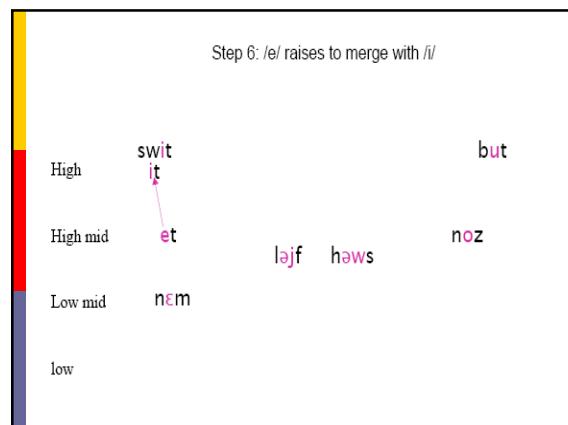
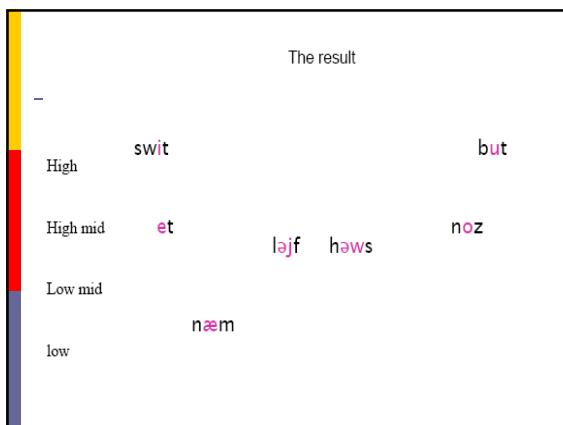
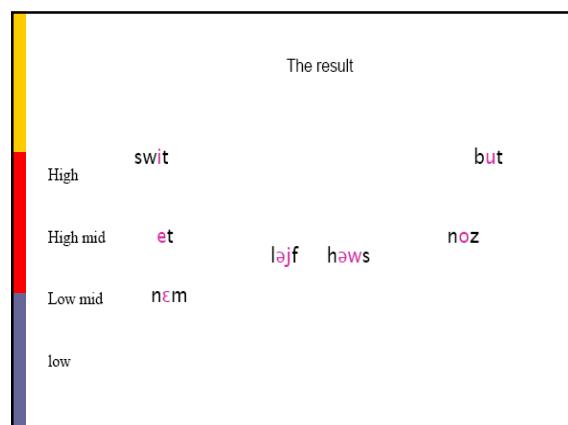
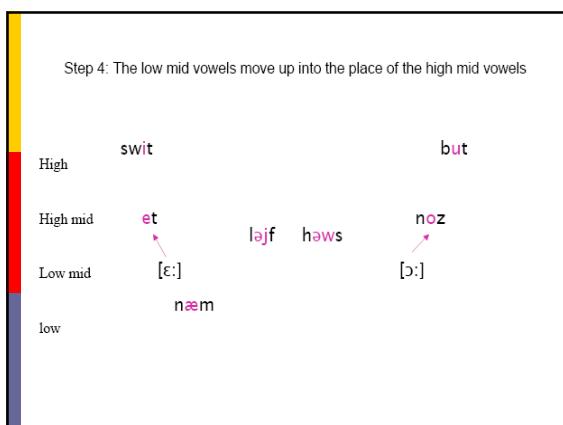
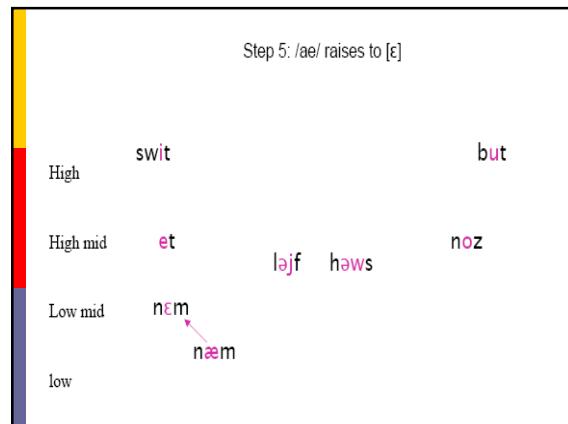
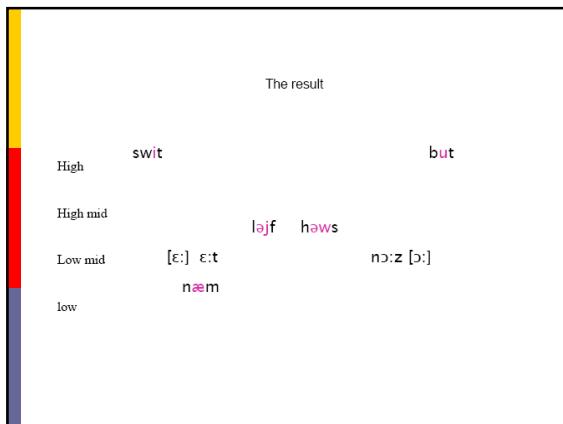
High	swit	but
High mid	ləf həws	
Low mid	[ɛ:] ē:t	nɔ:z [ɔ:]
low	[a:] nā:m	

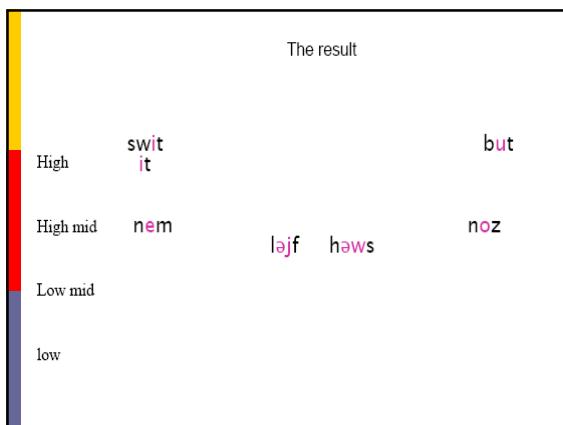
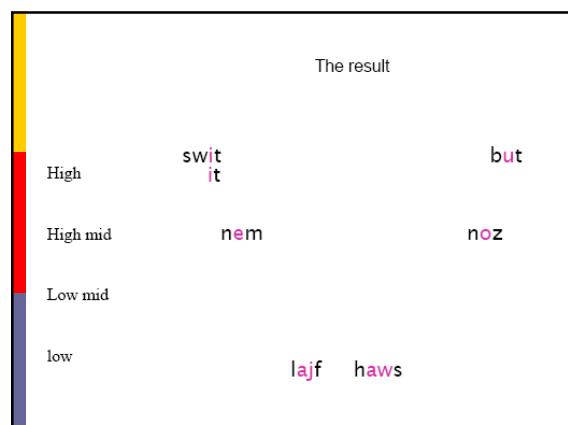
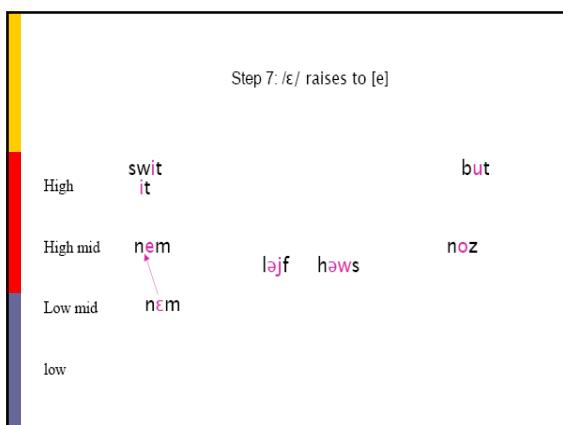
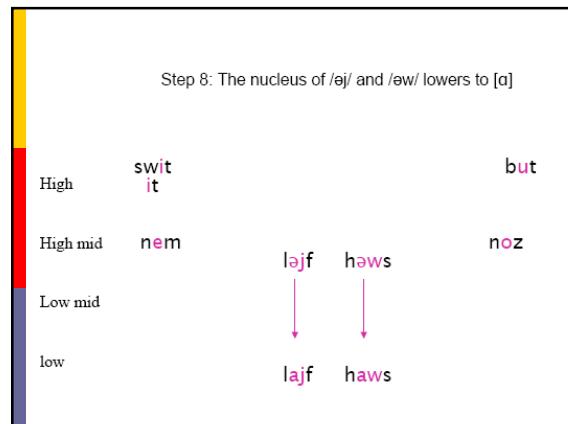
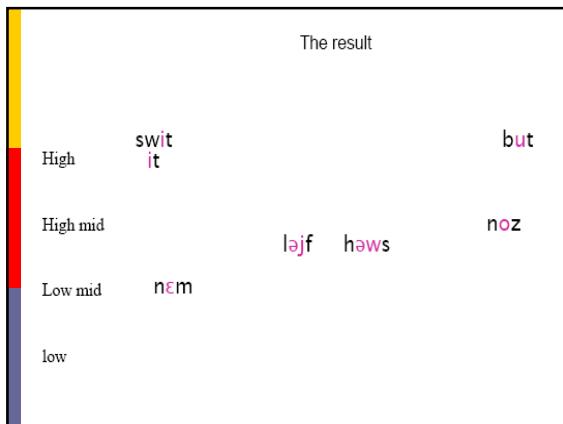
## The result

High		
High mid	[e:] swē:t	bo:t [o:]
Low mid	[ɛ:] ē:t	ləf həws
low	[a:] nā:m	nɔ:z [ɔ:]

## Step 3: /a/ raises and fronts to [æ]

High	swit	but
High mid	ləf həws	
Low mid	[ɛ:] ē:t	nɔ:z [ɔ:]
low	næm	[a:]





- Vowel length was closely tied to syllable structure in ME
- Some long vowels shortened in closed syllables
- Short vowels commonly lengthened in open syllables

- Since only long vowels participated in the Great Vowel Shift, we're left with some interesting alternations in modern-day English
- The vowel in OE **ke:p-te** shortened to **kep-te** so it didn't raise. While the vowel in OE **ce:pan** remained long and raised to /i/ hence today we have the present-past alternation **keep ~ kept**
  - OE **no-su** lengthened in ME to **no:se** so it raised to /o/. But **nostyrl** (< nosu + tyrel 'nose+hole') stayed short. Hence today we have **nose ~ nostril/nozzle**
  - OE **ste-lan** > ME **ste:le** hence **steal ~ stealth** etc.
  - **writ** < OE **writ**
  - **write** < OE **wri-tan**

## And this may also create minimal pairs

		V> $\tilde{V}$ /_N	N> $\emptyset$ /_C,##		
'big'	grand	grānd	grād	'rank'	grad
'good'	bon	bōn	bō	'beautiful'	bo

## Conditioned sound change

- begins with allophonic variation

		V> $\tilde{V}$ /_N
'big'	grand	grānd
'good'	bon	bōn

- Some sound changes are unconditioned. They apply to every occurrence of the phoneme, right across the lexicon.

*farina	>	harlo	'flour'	
*flore	>	hlū	'flower'	
*flakko	>	hlak	'flabby'	
*februario	>	hwre	'february'	
*filo	>	hiw	'string'	
*finestra	>	hjetro	'window'	
*folla	>	holo	'crazy'	
*furka	>	hurko	'pitchfork'	
*fundere	>	hum	'melt' inf.	
*formico	>	humigo	'ant'	
*flagello	>	hladefj	'thresher'	
*buffare	>	byha	'puff'	
*enfunare	>	éhurna	'put in the oven'	
*afatare	>	aharta	'stuff'	
*enflare	>	awhla	'swell up'	
*deforo	>	dehoro	'outside'	

Gascon \*f > h

## Conditioning environment changes

- Complementary distribution lost because of deletion

		V> $\tilde{V}$ /_N	N> $\emptyset$ /_C,##
'big'	grand	grānd	grād
'good'	bon	bōn	bō

- At this point, we have phonemic change.
- In this case, two new phonemes, /ā/ and /ō/ have emerged (actually, other nasalized vowels also emerged, but we're keeping it simple)

## Palatalization of Latin k

*karga	>	ʃarʒ	'charge, load up'
*karta	>	ʃart	'charter'
*kedru	>	sɛdr	'cedar'
*kentu	>	sã	'hundred'
*kippu	>	sɛp	'vine stock'
*korpu	>	kɔr	'body'
*kura	>	kyr	'cure'

$*k > \chi / \_ *a$		
<i>*blankare</i>	>	bläje 'bleach' inf.
<i>*kampo</i>	>	jä 'field'
<i>*kaptiare</i>	>	jase 'hunt' inf
<i>*kasteli:o</i>	>	jato 'castle'
<i>*kamisa</i>	>	jämiz 'shirt'
<i>*katta</i>	>	jat 'cat' f.
<i>*mänika</i>	>	mäj 'sleeve'
<i>*mekka</i>	>	mɛf 'wick, lock of hair'
<i>*kantare</i>	>	jäte 'sing' inf
<i>*kalore</i>	>	jaloer 'heat'
<i>*kamera</i>	>	jäbr 'room'

Latin was basically an SOV language  
But its case endings allowed speakers to mess with word order for stylistic and pragmatic purposes

<i>agricola canem videt</i>	'the farmer sees the dog'
<i>canem agriculta videt</i>	'the farmer sees the dog'
<i>agricola videt canem</i>	'the farmer sees the dog'
<i>videt agriculta canem</i>	'the farmer sees the dog'
<i>canis agricultam videt</i>	'the dog sees the farmer'
<i>agricolam canis videt</i>	'the dog sees the farmer'
<i>agricolam videt canis</i>	'the dog sees the farmer'
<i>videt canis agricultam</i>	'the dog sees the farmer'

## Sound-Syntax

- Sound change can have important consequences for morphology, and ultimately for syntax

## But...

$C > \emptyset / \_ \#$

*agricola cane vide*  
*cane agriculta vide*

Without the final consonants, the morphological distinction between nominative and accusative (subject and object) was lost.

So Latin's daughters had to replace morphology with fixed word order. As a result, the Romance languages are SVO languages.

## Latin singular noun paradigms

	'farmer'	'dog'
Nominative	<i>agricola</i>	<i>canis</i>
Genitive	<i>agricolae</i>	<i>canis</i>
Dative	<i>agricolae</i>	<i>cani</i>
Accusative	<i>agricolam</i>	<i>canem</i>
Ablative	<i>agricola</i>	<i>cane</i>

## Sound change is uniform '... and exceptionless': Part II

Consider:

OE *cnafa* [knava] > ModE [neɪv] 'knave'  
OE *cniht* [knixt] > ModE [najt] 'knight'

So what's the rule?

And what's the modern English reflex of OE *cyning* [kynɪŋ]?

## Sound change is uniform

### '... and exceptionless': Part III

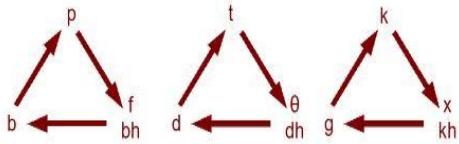
OE [kynɪŋ] > ModE [kɪŋ] 'king', of course.

Not [ɪŋ]; why?

Because the rule that gets rid of initial [k] only applies before [n].

The rule getting rid of initial [k] is still exceptionless, but we find that that's not exactly the rule: it doesn't get rid of initial [k] across the board, but only before [n].

## Grimm's law



## The Comparative Method

If we assume that sound-change is regular and exceptionless in this way, we can use systematic comparison of languages to see the relationships between them.

This is known as the **Comparative Method**.

## The Comparative Method: Part III

So far, so old news (but remember that this wasn't old news when the systematic correspondences were discovered in the nineteenth century!)

Is there anything that confirms the validity of the Comparative Method? Some language that shows a feature predicted by the Comparative Method, but wasn't known when the CM was worked out?

There is ...

## The Comparative Method: Part II

Sanskrit	Greek	Latin	Gothic	English	PIE
pita	pate:r	pate:r	fadar	father	*pəter-
padam	poda	pedem	fotu	foot	*ped-
bhratar	p'rare:r	frate:r	bropar	brother	*bhrater-
bharami	p'ero	fero	baira	bear	*bher-
sanah	hene:	senex	sinista	[senile]	*sen-
	tris	tre:s	þri	three	*trei-
	deka	dekem	taihun	ten	*dekm-
	he-katon	kentum	hund(rab)	hundred	*dkm-tom-

## The Comparative Method: Hittite

- Towards the end of the nineteenth century, there were a few places in IE reconstruction where the sound-changes didn't seem to be as regular as they should be. Relying on the **absolute regularity of sound-change**, Ferdinand de Saussure (in 1879) posited a set of *three* segments (based on the variable vowel quality in the daughter languages) that must have existed but had died out in all of the languages that were known up to then.
- He said that they were probably **laryngeals**.

## The Comparative Method: Hittite



- A little later, a set of clay tablets in familiar cuneiform script but in an unfamiliar language from ancient Turkey were deciphered (1600-1300BC); it was worked out that they were in fact in Hittite (an Anatolian language), which became the oldest IE language that we know.
- The texts did indeed contain segments in exactly the places that de Saussure had predicted. These days, we write them  $h_1$ ,  $h_2$  and  $h_3$ .

## Person marking in Latin verbs

	Sg.	Pl.
1st person	kanto	kantamos
2nd person	kantas	kantates
3rd person	kantat	kantant

'sing' present indicative

## The Comparative Method: limitations

- Note what the symbols  $h_1$ ,  $h_2$  and  $h_3$  look like.
- They don't imply anything about the pronunciation of the sounds, beyond the fact that they are probably laryngeal (as /h/ is).
  - That shows us that historical reconstruction is often like an exercise in logic, like the work on phonemes that we were doing earlier in the semester.
  - Historical reconstruction may not tell us anything about what PIE actually *sounded* like.

## Spanish dropped some of the consonants

- But maintained a differentiated paradigm

	Sg.	Pl.
1st person	kanto kanto	kantamos Kantamos
2nd person	kantas kantas	kantales kantaís
3rd person	kantat kanta	kantant kantan

## Sound-Change that isn't Regular

### Analogy

- The English ordinal number *sixth*, for example, goes back to an ancestral form containing the cluster *-kst-* (Latin cognate *sextus*), with a *-t-* that should not, according to the regular conditioning of Grimm's Law, have shifted to *th* after the fricative *-s-*. But the *-th* of the present-day English word has nothing to do with any failure of Grimm's Law. In fact, the Old English form was *sixta*, with *-t-*; the *-th* of *sixth* was introduced under the influence of the other ordinal numbers, where *-th* was phonologically regular (*fourth*, *seventh*, etc.).

## Gascon too

- But maintained differentiated paradigm

	Sg.	Pl.
1st person	kanti kanti	kantamos Kantamos
2nd person	kantas kantos	kantates kantats
3rd person	kantat kanto	kantant kanton

## But French?

- 4/6 forms identical!
- Consequences for French syntax?

	Sg.	Pl.
1st person	kanto jät	kantamos jätō
2nd person	kantas jät	kantates jäte
3rd person	kantat jät	kantant jät

## Morphological changes

- Backformation  
Worker ~ work, burglar ~ burgle, editor ~ edit, peddler ~ peddle
- Operate ~ operation, orientation ~ orientate (orient)
- Cherries ~ cherry
- Folk etymology  
Garter snake ~ garden snake
- Brydeguma (bride-man)
- Chaise longue ~ chaise lounge
- Nepotism ~ nephew-tism

## Only two identical forms

pronouns

	Sg.	Pl.
1st person	ȝə jät	nu jätō
2nd person	ty jät	vu jäte
3rd person	il/ɛl jät	il/ɛl jät

## Syntactic changes

- English, do-insertion for question and negation
- PIE had no articles (and neither do Latin or Sanskrit). Articles came later in Greek (not in Homer)
- Gender in Indo-Aryan

## Morphological changes

- Latin again, paradigm leveling

Nominative	honos	honos	honor
Genitive	honos-is	honor-is	honor-is
Accusative	honos-em	honor-em	honor-em

- Affected this paradigm but not *genus* or *navis*

## Jokes, only a linguist will tell

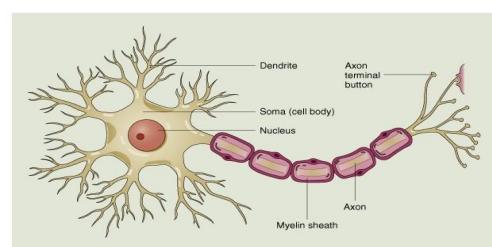
- Giles of Tottenham calls for ale at his favorite pub and is perplexed when the barmaid tells him that the fishmonger is next door.

## Neurolinguistics

Where brain meets mind



### Structure of the brain - neuron



© 2000 John Wiley & Sons, Inc.

### Evidence for “pre-wiring” of language

- **Universality:** language is universal to all humans and *only* humans
- **UG:** all human languages share certain universal features (UG)
- **Uniformity:** course of language acquisition is uniform (rate varies)
- **Learnability:** Certain aspects of language (UG) appear to be innate, not learned
- **Critical Period:** there is a critical period for language acquisition
- **Genetic:** There are genetic language disorders
- **Functional Modularity:** modular (selective) deficits (both acquired and developmental)
- **Structural Modularity:** The same brain regions are responsible for language in everyone

### Structure of the brain

- Weighs about 1300 to 1500 grams on average (newborn about 350-400g). Humans have the largest brain weight to body weight ratio among all earth's creatures.
- Contains over 100 billion ( $10^{11}$ ) brain cells, or *neurons* (and  $10^{14}$  synapses). Each neuron contains a soma (cell body), axon, and dendrites.
- Apart from their special structure, neurons are rather similar to normal body cells, except that they rarely ever divide. All the neurons you have now were generated by the age of 6 months.
- **Synapse:** junction between dendrites of one neuron and axon of another. In the cortex of the human brain, a typical neuron will have about 10,000 synapses.

### Neurolinguistics is...

Neurolinguistics tries to explore how the brain understands and produces language. This involves attempting to combine theory from neurology/neurophysiology (how the brain is structured and how it functions) with linguistic theory (how language is structured and how it functions).

Questions asked:

- Are we the only species with language/speech?
- Does our neurological make-up explain our linguistic ability – are there specialised regions?
- How did (evolutionary)/do (developmental) these brain mechanisms develop?
- Is speech special, or is language a part of general cognitive structures and abilities?

### Development of neurons

- All cells come from a single cell (fertilized egg). At first, all embryonic cells are undifferentiated. (After a neuron differentiates, it no longer divides.)
- **Migration:** After differentiation, neurons migrate to specific locations in the brain; arrangement is highly specific (months 1-4 in utero)
- **Branching:** Axons have a growth cone at the tip; guides placement of the axon; sensitive to chemical markers
- **Pruning:** Axon placement only gets to the right general neighborhood; fine tuned by pruning back inappropriate connections. “fire together, wire together”
- Development disorders can result from **incorrect migration** (lissencephaly (severe retardation), Down's syndrome, fetal alcohol syndrome) or **incorrect axonal connection** (cerebral palsy, autism, and other disorders)
- Synaptic concentration highest at about 2 yrs., and declines thereafter. Most pruning achieved before puberty. Adolescent/adult synaptic concentration about 60% of maximum.

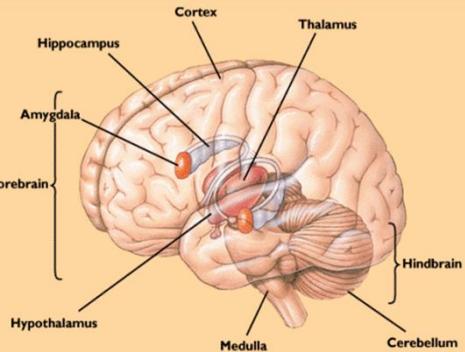
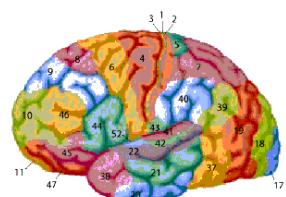
## Brain: The main components

- **Cortex/Cerebrum**, 2500cm<sup>2</sup>, 1.5-5mm thick, is responsible for most higher cognitive functions. This means that damage (even relatively shallow) to the brain can result in severe impairments in cognitive function
- Not everything happens in this cortical system; there are also subcortical connections: neural pathways that lead directly from one part of the cortex to another (like a "secret passage").

- **Cerebellum**, 'little brain', movement, balance, co-ordination
- **Medulla**, controls major autonomic functions (respiration, bp, heart rate) and **Pons**, relays information between the cerebellum and the cerebrum
- **Corpus callosum**, relays information between hemispheres
- **Spinal cord and the nerves** (peripheral nervous system)

## Brodmann's Areas (52)

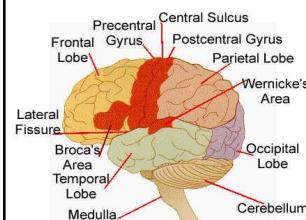
The cortex itself is structured, different parts of the cortex have different kinds of cell patterns (cytoarchitectonic fields or Brodmann's areas, often just known by their code numbers). Brodmann areas 1, 2 and 3 are the primary somatosensory cortex; area 4 is the primary motor cortex; area 17 is the primary visual cortex; and areas 41 and 42 correspond closely to primary auditory cortex.



- **Two hemispheres** connected by the **Corpus Callosum**: consists of ridges (**gyri**) and grooves (**sulci**). The hemispheres are *contralaterally*wired. All vertebrates have this characteristic.
- All invertebrates (e.g. prawns) have their brains wired to their body in a more obvious way; they also have their "spinal cord" on along the front and hearts in the back so perhaps the genetic mutation that led to backbones also caused the head to twist around 180 degrees?



## Brain: Landmarks



- **Central sulcus** (the big vertical groove)
- **Sylvian fissure or lateral fissure** (separates "wing")
- **Angular gyrus** bends around the tip of the SF
- There are four major lobes

## Four lobes



- **Frontal lobe:** From the front, to the central sulcus. Concerned with emotions, reasoning, planning, movement, (esp. LH) speech. Also involved in purposeful acts such as creativity, judgement, problem solving, and planning.
- **Temporal lobe:** The "wing" that is separated by the Sylvian fissure. Interpreting and processing auditory stimuli. Also involved in memory, meaning, language comprehension (esp. LH) and emotion.
- **Occipital lobe:** The back of the cerebrum; vision and ability to recognize objects.
- **Parietal lobe:** All the rest, between the frontal and occipital lobes; Connected with processing nerve impulses related to sensations of touch, pain, taste, pressure, and temperature. Also has language functions.

## Effect of experience on lateralization



- Lateralization occurs as child learns language
- In children with no linguistic exposure during childhood, no LH advantage for language
- Sensitive/critical period and lateralization (research of Helen Neville): LH dominant when using sign, but degree depended on age of learning sign: native signers were strongly lateralized; later learners were less strongly lateralized

## Plasticity, the ability to learn



- Ability for the brain to reorganize itself, esp. in response to stimuli, experience etc.
- Early development: experience important in forming proper neural connections
- Later learning: Doing more with less—existing neurons become more active. (Use of branching/pruning technique)
- Example: amputees and changing brain maps

## Problem of language



Linguistics	Neuroscience
<i>Fundamental elements of representation</i>	
• distinctive feature	?
• syllable	neuron
• morpheme	cell-assembly/ensemble
• noun phrase	population
• clause	cortical column
<i>Fundamental operations on primitives</i>	
• concatenation	long-term potentiation
• linearization	receptive field
• phrase-structure generation	oscillation
• semantic composition	synchronization

There is an absence of 'linking hypotheses' by which we explore how well-established brain mechanisms form the basis for linguistic computation.

## Plasticity and language learning



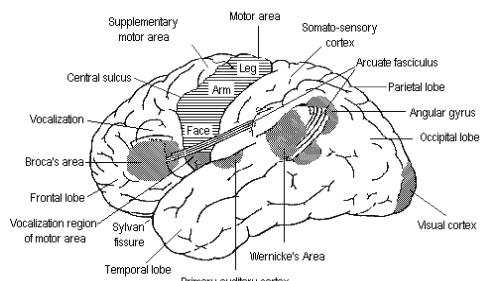
- Language is (for most of us) processed predominantly in the LH and trauma disrupts use of language
- 0-2 yrs: Damage to either hemisphere disrupts language. Language seems to be evenly distributed throughout the brain
  - 2-10 yrs: Lateralization begins to appear: language disrupted after LH damage only, but (near) complete recovery over time. Children with focal brain injuries involving language areas typically have little language impairment, and what language impairment they have tends to be very temporary. Adults are more likely to have permanent problems.
  - After puberty: plasticity seems to be lost; only limited recovery of language after LH damage
  - Hemispherectomies (removal of a hemisphere for medical reasons such as epilepsy), if performed in early childhood, many of the functions of the removed hemisphere are taken over by the remaining one; but this does not happen for adults.

## Function



- Certain areas of the brain are known to be responsible for *motor control* and for *sensory processing*
- These areas define a kind of map on the cortical surface, where different areas of the cortex correspond to different parts of the body
- Functions appear to be localized to certain areas of the brain, grossly identical across the species; thus we assume that similar localization must be true for human language as well

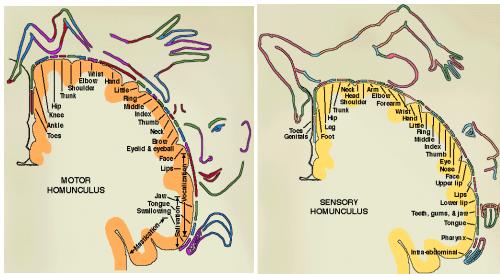
## Brain maps



### B. Neuroanatomical and neurobiological differences between LH and RH

- The sylvian fissure is longer and steeper on the LH—observable as early as 29<sup>th</sup> week of gestation
- The planum temporale (part of the auditory cortex/Wernicke's area) is larger in the LH than in the RH in 65/100 brains examined (and larger in RH in 10/100).
- Similar sylvian fissure asymmetries observed in some non-human primates (but not all); other asymmetries observed in non-primate species
- RH tends to mature earlier than LH; certain hormones (esp. testosterone) retard LH development

## Motor and Sensory Maps



## Studying the brain for language

### Single dissociation of function:

Find a group of patients able to perform some tasks but not others (e.g., count but not add, read but not write). This is suggestive that the two tasks are functionally and/or structurally modular. However, it could be that the two tasks are NOT modular, just that one is cognitively easier than the other (e.g., a patient who adds but can't do calculus, a patient who can play the scales, but not a musical piece).

## Lateralization

### A. Lateralization or asymmetric localization of a cognitive skill

- Left Hemisphere specialises in: Analytical skills (breaking things down), logic, language, complex motor sequences, temporal order
- Right Hemisphere specialises in: holistic skills ("big picture"), spatial cognition, music, non-verbal thought

### Double dissociation of function.

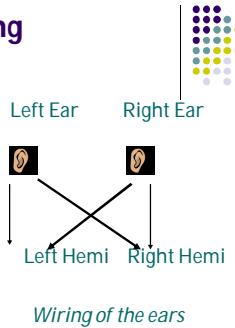
If there are two groups of patients who suffer from complementary disorders (say a and b), the two functions must be modular. It isn't just that one function is cognitively harder than the other because one group of patients can do a but not b and the other can do b but not a

- Broca's vs. Wernicke's Aphasia
- William's Syndrome (intact language) vs. Specific Language Impairment (impaired language)
- Asperger's Syndrome or high function autism (good linguistic competence, poor pragmatics) vs. Down Syndrome (poor linguistic competence, good pragmatics)

## Methods for studying lateralization-1

### Dichotic Listening Tasks:

- Contralateral connections: stronger than ipsilateral which are assumed to be blocked when contradictory information is presented to the ears
- Stimuli to left ear must undergo callosal relay to reach LH
- Dichotic presentation: One ear hears, say "cat", and the other hears "shoe." "What word did you hear?"
- % correct identification using right ear something like 7-9 percentage points higher than % correct identification using left ear (right ear advantage), e.g., 80% vs. 72%.
- Split-brain patient: very strong REA—about 90% correct for right ear, and 22% correct for left ear (similar for right hemispherectomy)
- Opposite pattern for identifying non-speech sounds



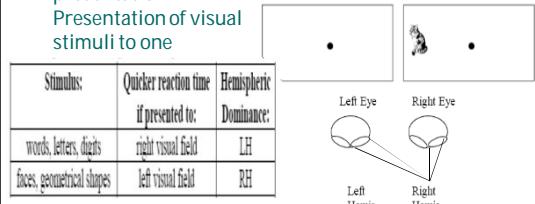
## Split brain studies

- The corpus callosum relays information between hemispheres. Split brain patients have had the connection severed surgically (usually because of intractable epilepsy). Auditory, haptic or visual stimuli can be presented to one or the other hemisphere. When language is presented in the LEFT ear, the split brain patient will report that they can hear sounds, but they can't understand it. (This is because their right hemisphere can HEAR the sounds, it just can't process the sound as linguistic.)

## Methods for studying lateralization-2

### Tachistoscopic presentation:

Presentation of visual stimuli to one



## Split brain studies

- Image shown to right visual field: object is named.
- Image shown to left visual field: object cannot be named, but can be pointed out, *using the left hand*.
- Chimeric images: image on the right will be named (man). Image on the left will be pointed out (woman).



## Methods for studying lateralization-3

WADA test (sodium amyta injections): Inject a (temporary) anesthetic to either the left carotid artery (feeds the left hemisphere) or the right carotid artery (feeds the right hemisphere). The injection causes temporary paralysis of the particular hemisphere.

The subject is first asked to do a linguistic task (telling a story, counting etc.), and then the injection is given either to the right or the left artery. If the person stops being able to do the task it suggests that language is lateralized to the 'paralysed' hemisphere. If they continue to do the linguistic task even after injection, that means that language is either bilaterally represented or is lateralized on the other (non-paralysed) hemisphere.

- Problem 1:** About 1% of the time, people have strokes. Thus, this method is only used for people who are neurosurgical patients. (It used to be done on healthy medical students and undergraduates.)
- Problem 2:** Patients have known brain disorders (usually intractable epilepsy), so it is a worry that they may have unusual lateralization anyway.

- **ERP (Event-related potential or evoked response potentials).** Change in an EEG related to a particular stimulus, factored out from the background brain activity. The ERP can be used as an on-line measure of "surprise" caused by ungrammatical sentences.
- Good news about this: can do with babies, children and adults. No risks whatsoever and pretty cheap. Very good temporal resolution (10-20 msec) but terrible spatial resolution (left vs. right hemisphere, front/middle/back).
- Wood (1975) English-speaking subjects were presented with speech sounds. When the task was to judge the pitch (not linguistic), the ERPs were more active on the right. When the task was to identify the syllables (linguistic), the ERPs were more active on the left.

## Methods to study neural functions -Open Brains

- **Electrocorticostimulation** (stimulate the brain' and see what it does): While a neurosurgical patient is AWAKE, stimulate regions of the brain and see what happens. Does the person 'hear voices'? Does the person stop being able to talk? Does the person spontaneously start talking.
  - **Electrocorticography** (record brain activity in response to exposure to linguistic/nonlinguistic material). While a neurosurgical patient is AWAKE, present linguistic/nonlinguistic material and see what regions of the brain are active.
- Problem: These patients have (usually) intractable epilepsy, so is this really NORMAL localization?

- **MEG:**
- magnetoencephalography**  
record surface magnetic fields; mapping brain activity by recording magnetic fields produced by electrical currents occurring naturally in the brain. This is like an ERP. Safe with good spatial and temporal resolution, but very expensive.



## Closed Brains

- **EEG, electroencephalogram :** Surface electrodes placed on scalp measure electrical potentials (firing neurons) in the brain areas under the electrode; provides a coarse measure of brain activity, compared to PET and MRI. Measurement is indirect, and so you can get false information from echoes, transduction along the skin, muscle movements, etc.

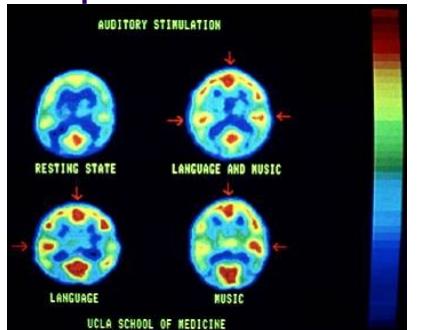


## Methods to study neural functions – PET Imaging

- **PET (Positron emission tomography).** Inhale/inject radioactive material – oxygen, fluorine. This mixes with the blood. The parts of the brain that have the most blood flow (greater metabolic activity) will be more radioactive than those with less blood flow. Regional cerebral blood flow (rCBF) is believed to correlate with brain activity. Provides very good spatial resolution, but scans are expensive, involve a hefty dose of radiation, provide terrible temporal resolution (approximately 60 secs), can study one task at a time, time consuming.



## PET picture



- **Tachistoscopic presentation** with native signers (research by Helen Neville). Fixation point, then number handshape. After fixation: must identify number first, then state the sign presented. Reaction time for identifying signs was quickest when presented to the right visual field, i.e., LH dominance for sign
- **Localization of visual/spatial information processing:** Native signers show LH dominance/superiority for dot localization while hearing subjects show RH superiority. Explanation: Since spatial localization is an important property of a signed language, this is localized in the LH (language hemisphere) in signers.



- **MRI and fMRI (Functional Magnetic Resonance Imaging, nuclear magnetic resonance imaging, NMR).** Put people in a very strong magnetic field and use the magnetic field to align the protons naturally present in the body. Then see how quickly the protons return to relaxed state. Haemoglobin can carry between 0 and 4 molecules of oxygen. The magnetic signature of Hb varies depending on how many oxygen molecules it carries. By measuring the amount of Hb that is deoxygenated, one has a view of how metabolically active the tissue is and which areas are co-opted for a specific task.

## Age of exposure and lateralization

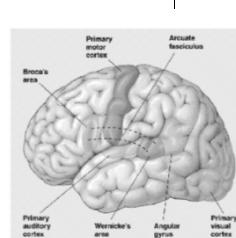
- EEG results for deaf signers: native signers showed strong lateralization (to LH), late learners showed weak lateralization (research by Helen Neville)
- Dichotic listening results for Genie: strong left ear advantage — 100% correct identification rate for left ear (RH superiority), 16% correct for right ear (research by Susan Curtiss)
- Tachistoscopic presentation results for Chelsea: no asymmetry was found for processing words: presentation to left and right visual fields showed about the same level of performance

## Sign and lateralization

- Sign language uses the visual-manual mode** (speech is aural-oral). Sign incorporates language structures, but relies heavily on visual/spatial relations—where is it processed?
- **EEG** (research by Helen Neville): LH more active when using sign, including “auditory” cortex. More recently, similar findings with PET and MRI.
  - **Sign aphasia:** Damage to LH causes disruptions in use of sign; damage to RH generally does not. Damage to RH disrupts spatial skills, *except* linguistic spatial relations such as correct use of signing space, correct movement of signs, etc.

## Language centres in the brain

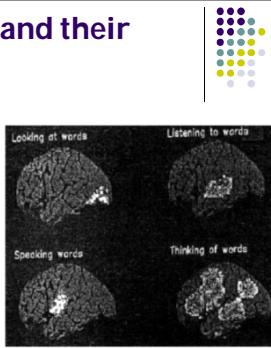
- Broca's area:** Located in L frontal lobe; adjacent to L motor cortex  
**Wernicke's area:** Located in L temporal lobe; adjacent to L auditory cortex  
**Arcuate fasciculus:** Nerve bundle connecting temporal and frontal lobes (there are two of them — the L *arcuate fasciculus* is important for language processing)  
**Angular gyrus:** Adjacent to Wernicke's area, near L visual cortex (above and to left: supramarginal gyrus)



## Language centres and their functions

Visual/Auditory Cortices: Receive incoming language stimuli

- **Broca's area:** Implementation of language into body movements (articulation); includes most aspects of grammars
- **Wernicke's area:** Comprehension of incoming speech/sign
- **Angular gyrus:** Reading, sound-spelling correspondences
- **Arcuate fasciculus:** Relays information from Broca's to Wernicke's areas



## Broca's contd.

- Lack of function words, suffixes; sometimes described as telegraphic speech ("agrammatism")
- Prone to "tip-of-the-tongue" phenomena; may produce the desired word if given the first syllable
- Usually shows good propositional content (impression: slow, but understandable) —prone to depression, frustration due to awareness of their speech difficulties
- Commonly accompanied by R weakness or paralysis
- Often show some recovery

## A-aphasia, without utterance - Broca's

"Lower Falls... Maine... Paper. Four hundred tons a day! And ah... sulphur machines, and ah...wood...Two weeks and eight hours...working...Working... working! Yes, and ah...sulphur. Sulphur and...Ah wood. Ah... handling!" (Goodglass 1973)

"Ah ... Monday ... ah, Dad and Paul Haney [himself] and Dad ... hospital. Two ... ah, doctors ... and ah ... thirty minutes ... and yes ... ah ... hospital. And er, Wednesday ... nine o'clock. And er Thursday, ten o'clock ... doctors. Two doctors ... and ah ... teeth. Yeah, ... fine."

## Aphasia – 2 Wernicke's

Examiner: What kind of work have you done?

Patient: We, the kids, all of us, and I, we were working for a long time in the ... you know ... it's the kind of space, I mean place rear to the spedawn ...

Examiner: Excuse me, but I wanted to know what work you have been doing.

Patient: If you had said that, we had said that, poomer, near the fortunafe, porpunate, tamppoo, all around the fourth of martz. Oh, I get all confused.

"Do you like it here in Kansas City?

"Yes, I am."

"I'd like to have you tell me something about your problem."

"Yes, I ugh can't hill all of my way. I can't talk all of the things I do, and part of the part I can go alright, but I can't tell from the other people. I usually most of my things. I know what can I talk and know what they are but I can't always come back even though I know they should be in, and I know should something eely..."

## Broca's contd.

- Relatively good comprehension of others
- Grammar-related difficulties in comprehension (ex: wh-movement, possessive's, etc.). Broca's aphasics seemed to have no trouble comprehending sentences like "The book that the girl is reading is yellow." When sentences are given that require more than just real-world knowledge, where syntax is crucial for understanding the meaning, they have much more trouble understanding. Thus they cannot distinguish between pragmatically reversible sentences: The cat chased the mouse vs. The mouse chased the cat
- Slow, dis-fluent speech ("motor aphasia," "expressive aphasia"); swearing, praying may be fluent

## Wernicke's contd.

- Poor comprehension; unable to follow simple commands ("receptive aphasia")
- Rapid, fluent speech that is usually devoid of meaning.
- Use of novel words, esp. nonsense words with real prefixes or suffixes appropriate to the grammar of the sentence: "I am *flimming*." ("jargon aphasia")
- Poor semantic judgements; which words are related 'mother, father, fish'.
- Often seem to be unaware that they have any language deficit.
- Not usually accompanied by other physical symptoms, poor recovery

## Basic picture

- The picture that emerged in light of these differences focused on production (impaired with Broca's aphasia) and comprehension (impaired in Wernicke's)
- As we will see, more detailed examinations reveal that Broca's aphasics have difficulties with comprehension also-- notable, when the syntax is 'complex', in ways which relate to what we have seen in the past



## Aphasia - Others

- Transcortical aphasias:** Lesion sites are not well understood
- transcortical motor aphasia**, Broca's area isolated— Able to repeat; poor comprehension and spontaneous speech; similar in some ways to Broca's aphasia (slow, agrammatic)
  - transcortical sensory aphasia**, Wernicke's area isolated— Impaired comprehension; good repetition and spontaneous speech; Repetition is mandatory (echolalia), no understanding of what is repeated
  - mixed transcortical aphasia or echolalia**, speech centres intact and disconnected—Only able to repeat; No comprehension or spontaneous speech; Correction of errors in repetition



## Aphasia - Others

### Global Aphasia

- Widespread linguistic deficit
- Damage to more extensive areas of LH
- May have R side weakness or paralysis



### Conduction Aphasia

- Involves damage to arcuate fasciculus (although cortical damage also present to produce "full" syndrome)
- Relatively good comprehension and spontaneous speech
- Unable to repeat what someone else has said
- Some phoneme errors in spontaneous speech; prob. due to lack of self-monitoring
- Difficulty reading aloud

**Anomia:** Small lesions in language areas. Good general language use except ability to find the appropriate word — an object may be described ('thing to eat', 'food that is not good for you' for 'cookie') rather than named, associated with posterior lesions

**Agraphia:** Inability to write words (phoneme to grapheme mapping is disturbed)



## Aphasia - Others

**Acquired Dyslexias:** Usually involve damage to angular gyrus and/or supramarginal gyrus

- **Deep dyslexia:** Inability to read non-words (*glurp*); Meaning extracted without being able to retrieve pronunciation (*poultry* might be read as *chicken* or *turkey*)
- **Surface dyslexia:** Only able to read phonetically; preserved ability to read non-words.
- **Angular gyrus:** In Japanese speakers, damage impairs ability to read kana (syllabic script) but not kanji (ideographic script); developmental dyslexics show abnormal angular gyrus activity and organization



## Aphasia - Others

**Pure Word Blindness or alexia:** Damage to the angular gyrus. Can speak, articulate but cannot read (grapheme to phoneme mapping). Able to spell words out aloud and to derive meanings from words spelled out. Can copy written words correctly. Loss of grapheme-phoneme rules – cannot read non-words or new words.

**Pure Word Deafness:** Damage to Heschl's gyrus near the auditory cortex. Cannot recognize spoken words, so comprehension and repetition is impaired. Can read and speak. Can 'hear' environmental sounds. Not deaf!



## Sign language and aphasia

- All reported cases of (acquired) sign aphasia result from LH damage
  - Signers with acquired aphasia show repeat classical aphasia typology:  
**posterior LH damage** (Wernicke's area): fluent, grammatical production, poor comprehension; anomia  
**anterior LH damage** (Broca's area): belabored, agrammatic sign; formation errors ; omission or misselection of affixes, good extra-linguistic communicative ability
- Important dissociations for processing of grammatical vs. non-grammatical visual-spatial relations (see next)



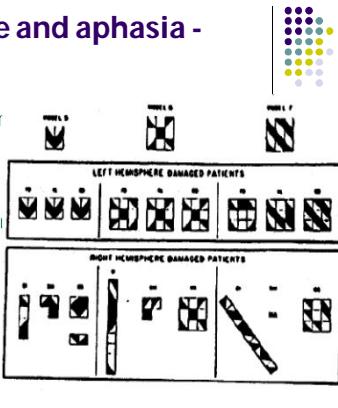
## Facial gestures and sign

- Affective facial gestures vs.
- Linguistic facial gestures



## Sign language and aphasia - Example

- LH damaged signer (top): Good performance on block design task
- RH damaged signer (bottom): Poor performance on block design task



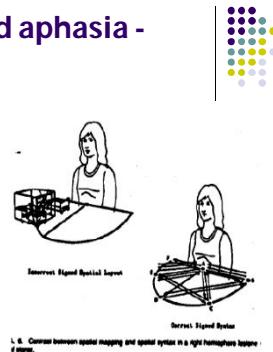
## Conclusion

- Language is highly specialized
- Our brains seem physiologically adapted for this function
- Language is double dissociated from intelligence (SLI and William's)
- Variation between brains is quite significant
- Don't know what happens exactly at the neuronal level
- but next – some studies on normal language



## Sign language and aphasia - Example

- RH damaged signer: Impaired non-grammatical use of space (top)
- Preserved grammatical use of space (bottom)  
 (In both examples, the patient was asked to describe her)



## Writing Systems

### The Place of Writing

- Writing is an addition in some sense; a famous quote:  
*Writing is not language, but merely a way of recording language by visible marks.*  
(Bloomfield, 1933)
- Writing systems are distinguished from other possible symbolic communication systems in that the reader must usually understand something of the associated spoken language to comprehend the text. In contrast, other possible symbolic systems such as information signs, painting, maps and mathematics often do not require prior knowledge of a spoken language.

### Socrates against writing in Plato's Phaedrus

- Writing, Socrates argues, is inhuman. It attempts to turn living thoughts dwelling in the human mind into mere objects in the physical world. By causing people to rely on what is written rather than what they are able to think, it weakens the powers of the mind and of memory. True knowledge can only emerge from a relationship between active human minds. And unlike a person, a text can't respond to a question; it will just keep saying the same thing over and over again, no matter how often it is refuted.

### But...

- Claiming that writing is not primary like spoken language is important; some additional points:
  - Writing is nevertheless an enormously important development in the history of human culture
  - Studying writing reveals much about some of the relatively abstract linguistic categories we have seen earlier in the course (particularly in phonology)
  - Studying reading also tell us about how linguistic information is processed by the brain

### Writing Systems: Context

- In the background: what we have looked at
  - Questions about whether SPOKEN language is an innate capacity of humans or a useful invention
  - Investigation of aspects of linguistic structure (sound, syntax, meaning)
  - How language is acquired by children: computed in the brain
- What is to come:
  - How languages vary across time (historical linguistics) in particular relationships

### Key Differences

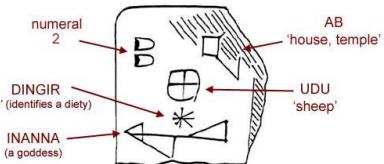
<u>Speech</u>	<u>Writing</u>
1) Present in all societies	Present in few, only recently
2) Learned before writing	Learned after speech, if at all
3) Acquired without instruction	Requires explicit instruction

### Basic components of the discussion

- Outline of the history of writing as currently understood, with reference to how early writing emerged
- Illustration of the distinct types of writing systems that are found, concentrating how elements of the writing system relate to different linguistic objects
- Objects of analysis:
  - What are the basic units of the writing system?
  - What linguistic objects do they pick out? Morphemes? Syllables? Phonemes?
  - How direct is the mapping between symbol and what it connects with?

### Example: Sumerian

- Early Development: A set of icons inscribed on clay, used for temple economic records:



**Meaning:** Two sheep received from temple of Inanna; or  
Two sheep delivered to temple of Inanna; or ...

### Origins of Writing



### Comments

- Note that only certain types of information— nouns, numbers, appear here
- Such tablets are difficult to interpret for this reason; they do not contain a full indication of grammatical structure and other aspects crucial to the spoken language
- Given certain conventions such tablets could be understood, but the system falls short of a full writing system

### Different Places/Times of Origin

- Four different times and cultures in which writing was invented (some question about contact in the Egyptian/Mesopotamian cases)
  - ◆ **Sumerian:** 3200 BCE (Cuneiform)
  - ◆ **Egyptian:** 3250-3050 BCE (?)
  - ◆ **Chinese:** 1200 BCE (?)
  - ◆ **Maya:** 292 CE
  - ◆ **Indus Valley:** 3500 BCE(?)



### Development

- Two principles in the developing system allow it to extend beyond the kind of notation seen above:
  - **Charades Principle:** Hint at the meaning or sound of a new symbol by making it look like an existing symbol
  - **Rebus Principle:** When a picture won't suffice (for e.g. an abstract concept or a suffix), use a symbol that is associated with a homophone (or something that sounds similar)

## Charades: Illustration

	Uruk ca. 3000	Lagash ca. 2400	Assyrian ca. 700
SAG 'head'			
KA 'mouth'			
GU 'eat'			
EME 'tongue'			

### Process

- 1) SAG 'head' begins as a drawing of a head
- 2) KA 'mouth' involves arbitrary marks at the position of the mouth
- 3) GU 'eat' adds a bowl to the previous

*Final form:* Note that (1-3) do not involve sound. The final example, EME 'tongue', does. This adds to KA 'mouth' a sign which has the pronunciation ME, giving: what relates to the mouth and sounds like [me]? (Subsequently the combination becomes conventionalized)

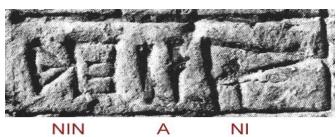
## What writing systems need

- at least one set of defined base elements/symbols/glyphs, individually termed **characters** and collectively called a **script**;
- at least one set of rules and conventions (orthography) understood and shared by a community, which arbitrarily pairs meaning to the base elements (graphemes), their ordering and relations to one another;
- at least one language (generally spoken) whose constructions are represented and recalled by the interpretation of these elements and rules;
- some physical means of distinctly representing the symbols by application to a permanent or semi-permanent medium, so they may be interpreted

## Rebus: Illustration



- The sign for A 'water' developed as shown above
- Sumerian also had a suffix [a] which meant 'in'
- The same sign for water was then used for this suffix
- In a more general development, the sign for water was then used for the phonetic value [a] more generally: so NIN-ANI 'his lady':



NIN  
'lady'  
A  
(‘water’)  
NI  
(‘oil’)

## Classification

- Traditional Terms:
  - **Ideographic:** Sign refers to an idea or general concept, not an actual word (in this sense not 'true' writing)
  - **Logographic:** Each sign refers to a specific word
  - **Syllabic:** Each sign refers to a syllable (e.g. CV)
  - **Alphabetic:** Each sign refers to a single sound (consonant or vowel)
- In practice these terms refer to symbols rather than systems, as systems may make use of more than one (example later)

## Another example



Bee

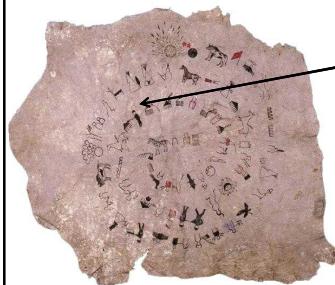


+  
Belief

Leaf

## Ideographic: Example

- Winter count, recording significant event, kept by Yankton band of Dakota Indians from 1800-1871



**Example:** Record of White soldiers attacking a fort



### Chinese ideographs

一	二	三	上	下	中	力	凸	凹
one	two	three	above	below	middle	strength	convex	concave (plough)

### Logographic, cont.

- Development of symbols:



Ma3 'horse'

Derived from:



Individual symbols in this system may be logographic, but for the majority of cases the signs use an indication of how the sign is pronounced.

### Egyptian ideographs

#### Egyptian Hieroglyphs:

*It is a complex system, writing figurative, symbolic, and phonetic all at once, in the same text, the same phrase, I would almost say in the same word.*

--Jean-François Champollion

	man, person
	woman
	people, mankind, Egyptians
	child, (to be) young
	old man, (to be) old, to lean
	man of authority
	noble person, dead person
	god, king
	king

### Extensions

- As noted on the last slide, not all signs in the Chinese system are logographic
- Many symbols have two parts
  - One indicates something about meaning
  - One indicates something about sound
- Example:



Ma1 'Mother'

### Logographic: Example

- Each sign stands for a word (maybe better, a morpheme) (From Greek "logos" = word, "graphos" = writing)
- Examples from Chinese; characters with the phonology of the words that they stand for (numbers indicate tones):

	Huo3 'fire'
	Mu4 'tree'
	人 man, person
	女 woman

目 eye, see   日 sun, day

### Example, Cont.

- Remember that we also saw:



Ma3 'horse'



Ma1 'Mother'

When we look at the way of writing Ma1 closely, we see:



mare: 母马 [mǔmǎ]

- The left side means 'female'
- The right side, the sign for 'horse', functions as a phonetic indicator

Sound meaning				
Semantic-phonetic compounds				
古 gǔ (person)	古 gǔ (to guess)	偏 biān (biased)	傲 ào (proud)	傍 bàng (beside)
言 gǔ (words)	詰 gǔ (commentaries)	論 pán (to quibble)	警 áo (to stander)	謗 bàng (to libel)
虫 gǔ (insect)	蛄 gǔ (mole cricket)	蝙 bān (bat)	螯 áo ([rab's] nippers)	螃 páng (crab)
金 gǔ (metal)	鉛 gǔ (cobalt)	鑊 áo (griddle)	鎊 bàng (pound sterling)	鏡 ráo (symbols)

## Syllabic: Example

- Syllabic systems employ symbols that stand for whole syllables; thus compare the following from Japanese hiragana with the English equivalent for the same pronunciation (meaning 'I' (polite form)):

わたくし  
wa ta ku shi

Ideographic - Logographic				
oracle bone	greater seal	lesser seal	clerkly script	standard script
rén (*níñ) human	人 rén	人 rén	人 rén	人 rén
nǚ (*nra?) woman	女 nǚ	女 nǚ	女 nǚ	女 nǚ
ěr (*nhǎ?) ear	耳 ěr	耳 ěr	耳 ěr	耳 ěr
mǎ (*mrǎ?) horse	馬 mǎ	馬 mǎ	馬 mǎ	馬 mǎ
yú (*ŋha?) fish	魚 yú	魚 yú	魚 yú	魚 yú
shān (*srāñ) mountain	山 shān	山 shān	山 shān	山 shān
ri (*niú) sun	日 ri	日 ri	日 ri	日 ri
yuè (*ŋŋot) moon	月 yuè	月 yuè	月 yuè	月 yuè
yǔ (*whā?) rain	雨 yǔ	雨 yǔ	雨 yǔ	雨 yǔ
yún (*wan) cloud	云 yún	云 yún	云 yún	云 yún

Hiragana										
	k	s	t	n	h	m	y	r	w	N
a	あ	か	さ	た	な	は	ま	や	ら	わん
i	い	き	し	ち	に	ひ	み	り		
u	う	く	す	つ	ぬ	ふ	む	ゆ	る	
e	え	け	せ	て	ね	へ	め	れ		
o	お	こ	そ	と	の	ほ	も	よ	ろ	を

**Notes:** The voiced obstruents g, z, d, b are indicated by adding two small strokes to the symbol for the corresponding voiceless consonant; and for historical reasons the /p/ symbols are derived from /h/ by adding a small circle. The N symbol is used when /n/ occurs in the coda of a syllable.

Hieroglyphs - Mayan		
<p>The writing system of the pre-Columbian Maya civilization of Mesoamerica, presently the only Mesoamerican writing system that has been substantially deciphered. The earliest inscriptions found which are identifiably Maya date to the 3rd century BCE in San Bartolo, Guatemala, and writing was in continuous use until shortly after the arrival of the Spanish conquistadors in the 16th century CE (and even later in isolated areas such as Tayasal).</p>		

Katakana											
片仮名 (カタカナ) katakana											
あ	ア	イ	イ	ウ	ウ	エ	エ	江	。	オ	於
ka	カ	キ	キ	ク	ク	ケ	ケ	介	コ	コ	己
sa	サ	シ	シ	ス	ス	セ	セ	世	ソ	ソ	曾
ta	タ	チ	チ	ツ	ツ	テ	テ	天	ト	ト	止
na	ナ	ニ	ニ	ヌ	ヌ	ネ	ネ	祢	ノ	ノ	乃
ha	ハ	ヒ	ヒ	フ	フ	ヘ	ヘ	部	ホ	ホ	保
ma	マ	ミ	ミ	ム	ム	メ	メ	女	モ	モ	毛
ya	ヤ	也	也	ユ	ユ	ヨ	ヨ	興			
ra	ラ	リ	利	ル	ル	レ	レ	礼	ロ	ロ	呂
wa	ワ	ヰ	ヰ	ヰ	ヰ	ヱ	ヱ	惠	ヲ	ヲ	乎
						ン	ン				无

## Mixed writing systems

- In some cases, a language's writing system viewed as a whole contains different types of objects; e.g. Japanese has
  - Kanji (symbols like the Chinese characters above)
  - Hiragana syllabic
  - Katakana syllabic (for foreign words etc.)

## Alphabetic

- Basic Principle:** each sign refers to a single sound
- This is of course familiar from English, although there are many deviations from the one-to-one pattern (Alphabetic-Consonantal etc.)
- In spite of the deviations, the underlying logic of the system is still alphabetic, as opposed to being syllabic or something else

## Kanji

- ◆ Basic repertoire of 1945 Chinese characters everyone has to learn
- ◆ About another 8,000 you might see in a novel but not in a government leaflet
- ◆ Many Chinese words have been borrowed into Japanese. Same characters get used for both
  - ◆ Kun yomi - Japanese reading
  - ◆ On yomi - Sino-Japanese reading

## Some exceptions

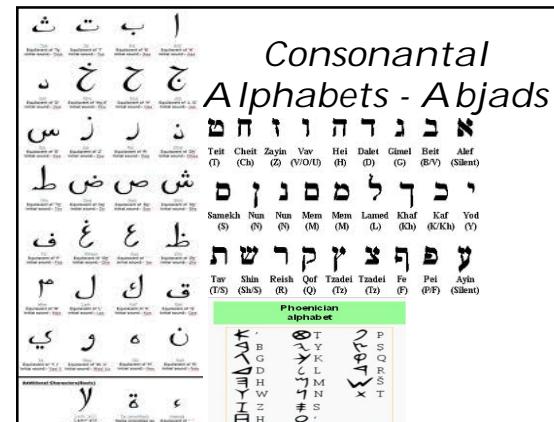
- Digraphs:** Two letters for one sound
  - Fish
  - See, too, saw
  - (and trigraphs: watch)
- Silent letters (i.e. present but do not contribute to the sound (gate, slight, cough))
- Same letters, different pronunciation (heat, great, bread)
- Different letters, same pronunciation (meat, meet, mete)

## For example

水	雨水	海水	水素	恐水病	水族館	水車
mizu	amamizu	kaisui	suiso	kyōsuibyō	suizokukan	suisha

(water element) (rainwater) (seawater) (hydrogen) (hydrophobia) (aquarium) (waterwheel)  
 (water vehicle)

Native Japanese: water = mizu  
 Sino-Japanese: water = sui



## Abugida/Aiphasyllabary

Transcription	śivō rakṣatu gīrvānabhiśārasāśvādatatparāṇ
Bali	શીવા રક્ષા પત્ર શીર્ષાંગામાં બાળ કુલાંગામાં માર્ગાં
Baybayin	સિદ્ધ દાન્સ વિન્ડ સિદ્ધ દાન્સ વિન્ડ
Bengali	শিবো রক্ষতু গীর্বানভি শারসাশ্বাদতপ্তপরাণ
Devanagari	शिवो रक्षतु गीर्वानभिशारसाश्वदतप्तपराण्
Grantha	ප්‍රිඩාමා පැකුතා තීව්‍යාලෙගාධාරාවාගාංජාත් කුංගාම
Gujarati	શિવો રક્ષતુ ગીર્વાનભિશારસાશ્વદતપ્તપરાણ
Gurmukhi	ਸਿਵੇ ਰਕਖਤੁ ਗੀਰਵਾਨਭਿਸ਼ਾਰਸਾਸ਼ਵਦਤਪਤਪਰਾਣ
Java	شیو بھا رکھتھو گیربھان بھارسادھے شوڈھے پاراھو
Kannada	ಶಿವೋ ರಕ್ಷತು ಗೀರ್ವಾನಭಿಶಾರಸಾಶ್ವದತಪ್ತಪರಾಣ
Malayalam	ശിവോ രക്ഷതു ഗീര്വാനഭിഷാരസാശ്വദതപ്തപരാണ്
Oriya	ଶିବୋ ରକ୍ଷତୁ ଗୀର୍ବାନଭିଶାରସାଶ୍ଵଦତପ୍ତପରାଣ
Sinhala	ශිවේ රක්ෂා ගීර්බානභි සාරසා එංජා පෙන් පෙන්
Tamil	ஷிவோ ரக்ஷது கீர்வானப்பாஷாலாஸ்வாததப்பாராங்
Telugu	శివో రక్షతు గీర్బానభిశారసాశ్వదతప్తపరాన
Tibetan	ཤི ཤ ས ལ ཡ ས ཤ ས ཤ ས ཤ ས ཤ ས ཤ ས ཤ ས
Thai	ສිවො ຮක්ෂතු ກීර්බානභි ສාරසා ສ්වද ຕප්තපරාນ

## Linear B logograms



## Cherokee Syllabary

a	e	i	o	u	v [ə]
D a	R e	T i	ɸ o	ɔ' u	i v
ʂ ga ɸ ka	f ge	y gi	A go	J gu	E gv
χ̄ ha	ɸ he	ɸ hi	F ho	Γ hu	χ̄ hv
W la	d̄ le	P li	G lo	M lu	ɿ lv
ɸ̄ ma	ɒ̄ me	H mi	ɔ̄ mo	ɟ̄ mu	
θ̄ na t̄ hna G nah	ʌ̄ ne	h ni	Z no	q̄ nu	θ̄ nv
ɿ̄ qua	ɔ̄ que	ɸ̄ qui	V̄ quo	ω̄ quo	ɛ̄ quv
ḡ s U sa	ɹ̄ se	b si	ɸ̄ so	ɣ̄ su	R sv
l̄ da W ta	S de ɿ̄ te	ɿ̄ di ɿ̄ ti	A do	S du	θ̄ dv
ʂ̄ dla ɿ̄ tla	L tle	C tli	ɸ̄ tlo	θ̄ tlu	P tlv
Ḡ tsa	V̄ tse	Ir̄ tsi	K̄ iso	ɟ̄ tsu	C tsv
Ḡ wa	ɹ̄ we	θ̄ wi	θ̄ wo	ḡ wu	6 vv
χ̄ ya	β̄ ye	ɸ̄ yi	ɸ̄ yo	Ḡ yu	B vv

Invented by a Cherokee silversmith called Sequoyah in the 1820s

## Undeciphered Scripts

- The Phaistos Disk was found in the Minoan Palace of Phaistos on Crete in 1908 and is thought to date from the 17th century BC. On it is inscribed an unknown script using pre-fabricated seals.



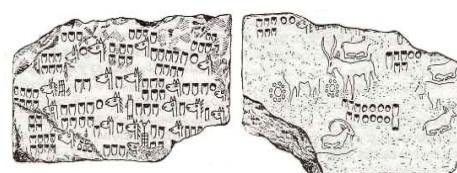
## Linear B Syllabary Mycenaean

ㅏ	ㅓ	ㅗ	ㅜ	ㅡ	ㅣ	ㅑ	ㅕ	ㅘ	ㅙ	ㅛ
a	da	ja	ka	ma	na	pa	qa	ra	sa	ta
ㅐ	ㅔ	ㅖ	ㅖ	ㅖ	ㅖ	ㅖ	ㅖ	ㅖ	ㅖ	ㅖ
ㅓ	ㅓ	ㅓ	ㅓ	ㅓ	ㅓ	ㅓ	ㅓ	ㅓ	ㅓ	ㅓ
ㅗ	ㅗ	ㅗ	ㅗ	ㅗ	ㅗ	ㅗ	ㅗ	ㅗ	ㅗ	ㅗ
ㅜ	ㅜ	ㅜ	ㅜ	ㅜ	ㅜ	ㅜ	ㅜ	ㅜ	ㅜ	ㅜ
ㅡ	ㅡ	ㅡ	ㅡ	ㅡ	ㅡ	ㅡ	ㅡ	ㅡ	ㅡ	ㅡ
ㅣ	ㅣ	ㅣ	ㅣ	ㅣ	ㅣ	ㅣ	ㅣ	ㅣ	ㅣ	ㅣ
ㅑ	ㅑ	ㅑ	ㅑ	ㅑ	ㅑ	ㅑ	ㅑ	ㅑ	ㅑ	ㅑ
ㅕ	ㅕ	ㅕ	ㅕ	ㅕ	ㅕ	ㅕ	ㅕ	ㅕ	ㅕ	ㅕ
ㅘ	ㅘ	ㅘ	ㅘ	ㅘ	ㅘ	ㅘ	ㅘ	ㅘ	ㅘ	ㅘ
ㅙ	ㅙ	ㅙ	ㅙ	ㅙ	ㅙ	ㅙ	ㅙ	ㅙ	ㅙ	ㅙ
ㅛ	ㅛ	ㅛ	ㅛ	ㅛ	ㅛ	ㅛ	ㅛ	ㅛ	ㅛ	ㅛ

(early form of Greek)

## Proto Elamite

first appeared in about 3100 BC in Suse (Susa), the capital of Elam, in south-western Persia (modern Iran). The Proto-Elamite script is thought to have been developed from an early Sumerian script and consists of about 1,000 signs and is therefore thought to be partly logographic. It has yet to be deciphered, and the language it represents is unknown.

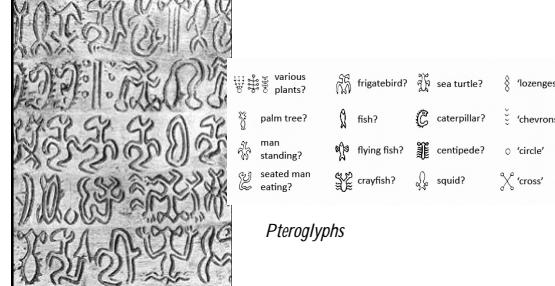


## Old Elamite

- Old Elamite was a syllabic script derived from Proto-Elamite and was used between about 2250 and 2220 BC, though was probably invented at an earlier date. Old Elamite has only been partially deciphered, mainly by Walter Hinz.*
- Old Elamite consisted of about 80 symbols and was written in vertical columns running from top to bottom and left to right.*



## Rongorongo - Easter Island



## Indus Valley/ Harappa

The term **Indus script** (also **Harappan script**) refers to short strings of symbols in use during the Early Harappan and Mature Harappan period, between the 3500-2000BC.

The first publication of a Harappan seal dates to 1873, in a drawing by Alexander Cunningham. Since then, over 4000 symbol-bearing objects have been discovered, some as far afield as Mesopotamia. In the early 1970s, Iravatham Mahadevan published a corpus and concordance of Indus writing listing about 3700 seals and about 417 distinct signs in specific patterns. The average inscription contains five signs, and the longest inscription is only 17 signs long. He also established the direction of writing as **right to left**.



## Indian Languages and Writing

PROTO-SINAITIC SCRIPT 19 c. BCE	
<u>Ugaritic</u> 15 c. BCE	<u>Proto-Canaanite</u> 14 c. BCE
<u>Phoenician</u> 12 c. BCE	
Greek 8 c. BCE	Aramaic 8 c. BCE
Etruscan 8 c. BCE	<i>Kharosthi</i> 6 c. BCE
Latin 7 c. BCE	<i>Brāhmaī</i> & Indic 6 c. BCE
Runic 2 c. CE	<i>Bhattiprolu</i> script
Coptic 3 c. CE	<i>Telugu</i> script
Gothic 3 c. CE	<i>Brahmic abugidas</i>
Armenian c. 405	Oriya 10 c. CE
Georgian 5 c. CE	Bengali 11 c. CE
<i>Glagolitic</i> c. 862	Devanagari 13 c. CE
Cyrillic c. 940	Hebrew 5c BCE
	Thaana 4c. BCE
	Pahlavi 5c. BCE
	Avestan 4c. BCE

## Mesoamerican Scripts

	Maya	Zapotec	epi-Olmecc	Aztec	Mixtec
snake					
jaguar					
hill					
moon					
human/deity head					
hand compound					
quadripartite shape					

Possible derivation of Brahmi from the Phoenician script														
Greek	Α	Β	Γ	Δ	Ε	Υ	Ζ	Η	Θ	Ι	Κ	Λ	Μ	Ν
Phoenician	𐤀	𐤁	𐤂	𐤃	𐤄	𐤅	𐤆	𐤇	𐤈	𐤉	𐤊	߱	߳	ߴ
Aramaic	܀	܁	܂	܃	܄	܅	܆	܇	܈	܉	܊	܋	܌	܍
Brahmi	ା	ି	ି	ି	ି	ି	ି	ି	ି	ି	ି	ି	ି	ି
Devanagari	अ	ब	ग	ध	त	द	थ	ठ	क	क्ष	म	न	ण	া
Tamil	அ	ப	க	த	உ	ஈ	ஞ	ஞ	ஞ	ஞ	ஞ	ஞ	ஞ	ஞ
Kannada	ಾ	ಿ	ಿ	ಿ	ಿ	ಿ	ಿ	ಿ	ಿ	ಿ	ಿ	ಿ	ಿ	ಿ
IAST	a	b	g	d	t	dh	th	th	k	kṣ	m	n	ṇ	া

## Kharosthi



- The **Kharoṣṭī script**, is an ancient abugida (or "alphasyllabary") used by the Gandhara culture ancient South Asia to write the Gāndhārī and Sanskrit languages (Tocharian and some Prakrits also). It was in use from the middle of the 3rd century BCE until it died out in its homeland around the 3rd century CE.
- It was also in use along the Silk Road where there is some evidence it may have survived until the 7th century.

## Brahmi



- The modern name given to the oldest members of the Brahmic family of scripts. The best-known Brahmi inscriptions are the rock-cut edicts of Ashoka in north-central India, dated to the 3rd century BCE. These are traditionally considered to be early known examples of Brahmi writing.
- Recent discoveries have revealed earlier epigraphy in Tamil-Brahmi, a Southern Brahmic alphabet found on pottery in South India and Sri Lanka dating from before the 6th century BCE Sangam period. Southern Brahmi gave rise to Tamil Brahmi, Vatteluttu and Pallava Grantha scripts that diversified into many South East Asian scripts like the Mon script in Burma, the Javanese script in Indonesia and the Khmer script in Cambodia.
- Northern Brahmi gave rise to the Gupta script during the Gupta period, which in turn diversified into a number of cursive forms during the Middle Ages, including Siddham, Sharada and Nagari. The script was deciphered in 1837 by James Prinsep, an archaeologist, philologist, and official of the British East India Company. Brahmi was an abugida.
- Brahmi was ancestral to most of the scripts of South Asia and Southeast Asia, several Central Asian scripts such as Tibetan and Khotanese, and possibly, in part, Korean Hangul.

- The earliest likely contact of the Hindu Kush region with the Aramaic script occurred in the 6th century BCE with the expansion of the Achaemenid Empire under Darius the Great to the Indus valley. It appears that no use of any script to write an Indo-Aryan language occurred before the reign of Emperor Ashoka in the 3rd century BCE, despite the evident example of Aramaic.
- Megasthenes, an ambassador to the Mauryan court only a quarter century before Ashoka, noted explicitly that the Indians "have no knowledge of written letters" (*Indica*). This might be explained by the cultural importance at the time (and indeed to some extent today) of oral literature for history and Hindu scripture.

## Sound and Syntax in the brain

Neuroimaging

### Basics of fMRI

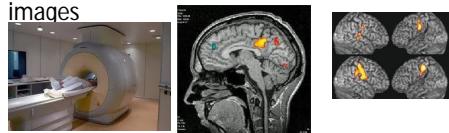
- Uses extremely strong magnetic field
- Measures changes in local blood flow/blood oxygenation
- Excellent spatial resolution
- Hemodynamic (= blood-related) response is slow in comparison with neuronal activation.

### Syntax in the brain

- As we saw in the last lecture, it has been proposed that 'syntax' resides in Broca's area
- This finding connects in some ways with what is found in the aphasia literature, although the situation is quite complicated
- Today: Introduce functional Magnetic Resonance Imaging (fMRI), and look at a few experiments that involve syntax

### Procedure

- Subjects lie within the bore of the magnet, with heads fixed
- Experimenters acquire structural images of the subjects' brains
- Functional images, showing changes in blood flow/oxygenation, are acquired and registered (composed together) with the structural images



### Spatial and Temporal Dimensions

Imaging techniques differ in spatial and temporal sensitivity.

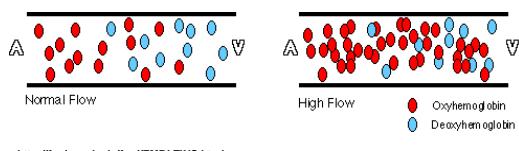
<u>Technique</u>	<u>Spatial</u>	<u>Temporal</u>
ERP	10-20mm	1msec
MEG	5-10mm	1msec
PET	<5mm	30sec
fMRI	<2mm	~1sec

### Further Details

- During changes in neuronal activity, there are local changes in the amount of oxygen in the tissue
- The changes in blood oxygenation are detected by the scanner
- The result is an MRI signal called BOLD: Blood-Oxygen-Level-Dependent

## Cortical Activity and Blood-Flow

Cortical activation results in a local increase in oxygenated blood, and blood flow, without an increase in oxygen consumption.



<http://fmri.ucsd.edu/fmri/FMRI-TINS.html>

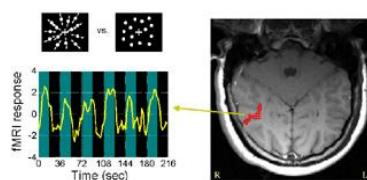
## Basics

- **Question posed:** Do computationally more demanding sentences result in more brain activation than less complex sentences?
- Focusing on areas known to be involved in language, i.e. Wernicke's and Broca's

## Illustration

Areas in which neuronal activity occurs show patterns of activation which the scanner detects.

The image is paired with a graph showing the blood flow in that region by experimental condition (moving vs. stationary dots).



## Stimuli

- Three types of sentences, with the same number of words:
  - **Active:** *The reporter attacked the senator and admitted the error.*
  - **Subject relative clause:** *The reporter that t attacked the senator admitted the error.*
  - **Object relative clause:** *The reporter that the senator attacked t admitted the error.*

## Study 1

- Just, Carpenter, Keller, Eddy, and Thulborn (1996), *Science* 274
- Cortical activation and sentential complexity

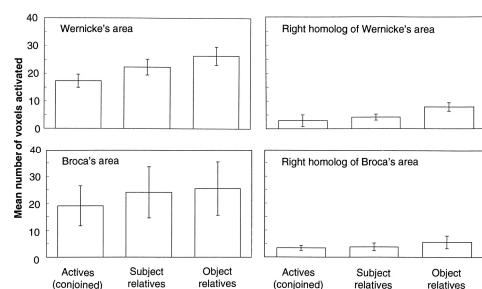
## Control

- Something to control for some of the visual components of the sentence-reading part
- Consonant strings:
  - Pws ntkggrfm zjkjrng kwtdc sbfght swn mrjbxq kgt mxbtq

## Design and Task

- Sentences were arranged in sets of four to five of the same type
- Following each sentence, subjects answered 'True' or 'False' to a comprehension probe:
  - "The reporter attacked the senator, True or False?"

## Cortical Activation



## Subject Performance

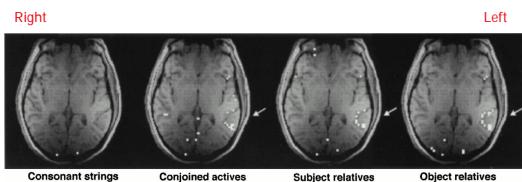
- Reaction time and error rate data were collected for most of the subjects
- Analysis of the behavioural data showed these indices of complexity to increase monotonically in the expected way with the increasingly complex sentence-types

## Interpretation

- It is suggested that Wernicke's is involved in something other than purely lexical matters (since these are constant across conditions)
- Broca's area is involved in something as well (perhaps syntax; it's hard to tell)
- Right homologues are implicated in processing as well
- **Overall:** The brain's response to increased processing demands is to recruit more tissue in each area in a network of cortical areas

## Images

The analysis looked at the number of activated voxels (volumetric picture element) per condition: more of these are present from left to right.



NOTE: Left is Right and Right is Left in these images

## Study 2

- Dapretto and Bookheimer (1999), *Neuron* 24
- fMRI study, attempting to dissociate syntax and semantics in sentence comprehension

## Motivation

- A number of prior studies showing activation in Broca's area during tasks of *different* syntactic complexity
- The idea that this does not establish anything about exclusivity— perhaps Broca's area would also be modulated by increasing demands along another, non-syntactic dimension

## Examples: Semantics

- Same
  - The **lawyer** questioned the witness.
  - The **attorney** questioned the witness.
- Different
  - The man was attacked by the **doberman**.
  - The man was attacked by the **pitbull**.

## Design Basics

- Auditory presentation of pairs of sentences
- Subjects had to decide whether or not the meaning of the two sentences differed
- Sentence pairs were classified as **semantic** or **syntactic**; in each case, there were 'same' and 'different' types

## Examples: Syntax

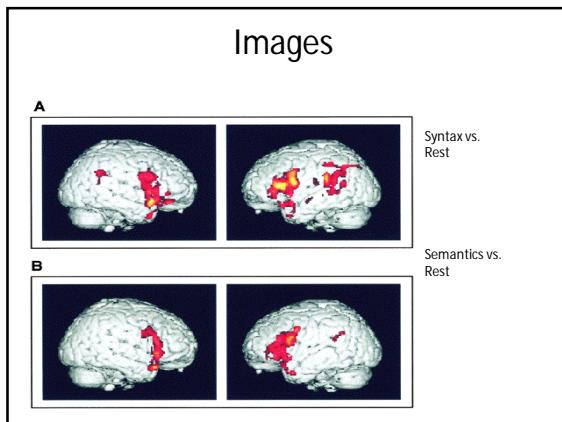
- Same (active-passive)
  - The policeman arrested the thief.
  - The thief was arrested by the policeman.
- Different
  - The teacher was outsmarted by the student.
  - The teacher outsmarted the student.

## Conditions

- **SEMANTIC:** Pairs of identical sentences, in which one word was replaced either by a synonym or a different word.
- **SYNTACTIC:** Sentences in the pair were in different voices (active vs. passive) or had word-order differences

## Justification of Stimuli

- In the semantic condition: Judgement is supposed to rely on single word meanings
- In the syntactic condition: Judgement requires computation and comparison of two syntactic structures
- Authors note that both kinds of processing are present in each task, and expect the two conditions to differ in relative 'weight' of syntax and semantics



### Interpretation

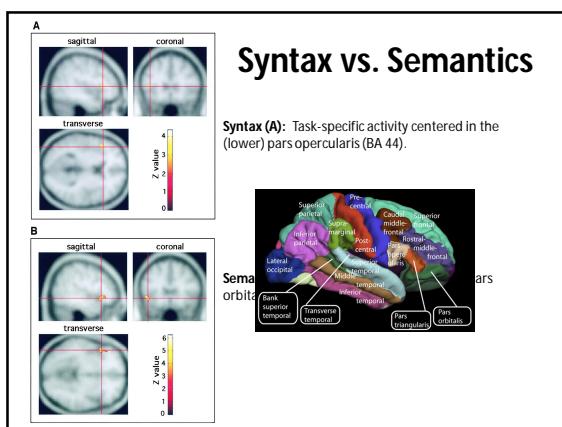
- Authors
  - Syntax: BA 44 is particularly involved in syntactic processing
  - Semantics: BA 47 is selectively involved in the processing of lexico-semantic information
- Spend some time thinking about how strong this kind of conclusion could be, given what happened in the experiment etc.

### Results (General)

- Many areas associated with language are found to be active in this experiment
- What about the syntax vs. semantics comparison?

### Wrapping up syntax

- Broca's area (and other language areas) can be modulated in different tasks
- How this connects with the idea of specialization is complicated
- Functional neuroimaging is impressive technologically, but the connections with cognitive hypotheses are in the nascent stages



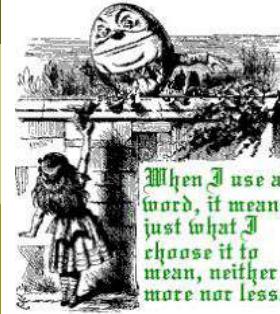
## Semantics: The meaning of meaning

### Two main areas of semantics

- The meaning of words (lexical semantics)
- The meaning of sentences (sentential semantics)

Speakers agree upon meanings and are not free to change or redefine it like Humpty Dumpty!

- A single semantic property may be a part of many words:
  - tigress, actress, doe, mare, girl, maiden, debutante, vixen, ewe, cow, nurse (?) **female**
  - doctor, dean, bachelor, professor, child, baby, parent **human**
  - child, baby, infant, youth, adolescent **youth**
  - father, uncle, bachelor, man **male**
  - darken, kill, uglify, feed, sink **cause**
  - walk, run, plod, canter, trot, glide, skip **move**
  - hit, kiss, touch, push, pinch, kicks **contact**
  - smell, hear, feel, touch, taste **sense**
  - build, imagine, make, design **create**



*When I use a word, it means just what I choose it to mean, neither more nor less.*

'And only *one* for birthday presents, you know. There's glory for you!' 'I don't know what you mean by "glory".' Alice said.  
Humpty Dumpty smiled contemptuously. 'Of course you don't -- till I tell you. I meant "there's a nice knock-down argument for you!"'  
'But "glory" doesn't mean "a nice knock-down argument".' Alice objected.  
'When *I* use a word,' Humpty Dumpty said, in rather a scornful tone, 'it means just what I choose it to mean -- neither more nor less.'  
'The question is,' said Alice, 'whether you *can* make words mean so many different things.'

### Semantic Properties

- Pieces of information on which speakers of a language agree
- ```
graph TD; assassin([The assassin killed Mr. Humphrey]); assassin --> murderer[murderer, human]; Humphrey[Mr. Humphrey] --> important[Important person]
```

The diagram consists of two ovals. The left oval contains the sentence "The assassin killed Mr. Humphrey". Two arrows point from this sentence to two separate labels below: "murderer, human" on the left and "Important person" on the right.

- Additional semantic properties make for the finer distinctions in meaning:
  - move:** walk, run, plod, trot, glide, stalk etc.
  - slow:** plod
  - menace:** stalk

## Evidence for semantic properties

- Speech errors or slips of the tongue
  - bridge of the nose/neck
  - when my gums/tongues bled
  - he came too late/early
  - Mary was young/early
  - that's a horse of another colour/race
  - he has to pay her alimony/rent
  - the lady with the dachshund/Volkswagen

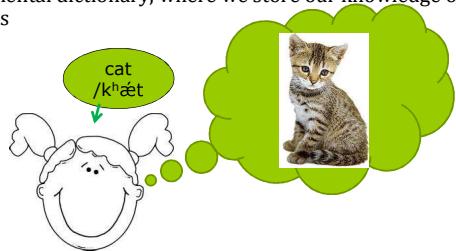
"The law I sign today directs new funds... to the task of collecting vital intelligence... on weapons of mass production."

## From Fromkin

- (29) (a) My data consists [mownlij] – [mejstlij] ... (mainly/mostly)
  - (b) I swindged [swɪndžd] ... (switch/changed)
  - (c) It's a lot of [ba] – [braðl] (bother/trouble)
  - (d) She's a real [swɪp] chick (swinging/hip)
  - (e) it's a [spajrətv] (spirant/fricative)
  - (f) a tennis [æθlər] (player/athlete)
- (30) (a) I really like to – hate to get up in the morning
  - (b) It's at the bottom – I mean – top of the stack of books
  - (c) This room is too damn hot – cold
  - (d) the oral – written part of the exam

## Lexicon

- our mental dictionary, where we store our knowledge of words



- Variation in space (regional differences): braces, suspenders
- Variation in time (historical change) – hound

## Lexicon and semantic properties

- Lexicon stores amongst other information, the semantic properties of morphemes
- Words that share a semantic property are part of a semantic class
- And semantic classes may intersect
  - Female, young > girl (+human), filly (+equine)
- Nouns: count vs. mass, abstract vs. concrete
- If something swims, it is in a liquid; if something is splashed it must be liquid  
*I saw a gnarle swimming in glish*

Glish can co-occur with *pour, drink, leak*; there may be *droplets of glish* but one cannot *saw* or *bend glish*.

## Languages and semantic properties - variation

- Navajo:
  - leh [+long, +flexible], example *rope*
  - tuh [[+long, -flexible], example *spear*
- Swahili
  - m- (sg), wa- (pl): mtoto *child* watoto *children* (human)
  - ki- (sg), vi- (pl): kiti *chair* viti, *chairs* (artefacts)

## Sense and reference

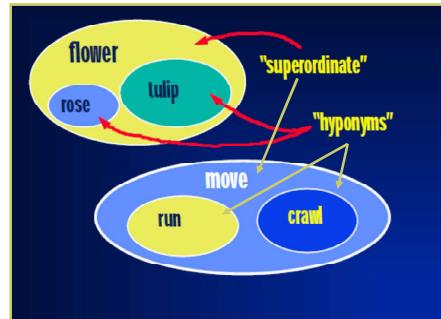
|                                  |                                          |
|----------------------------------|------------------------------------------|
| □ my father                      |                                          |
| □ my mother's husband            | Same referent<br>Different sense         |
| □ the best fisherman that I know |                                          |
| □ my best friend                 |                                          |
| □ the president                  | Different referent in different contexts |
| □ my brother                     |                                          |
| □ the present King of England    | No referent                              |
| □ George Bush                    |                                          |
| □ Hosni Mubarak                  |                                          |

## Lexical relations: Relatedness between words

### Hyponymy - Hypernymy

Relationship between more general words and specific instances of it

- |                             |                     |
|-----------------------------|---------------------|
| ❑ square                    | rectangle           |
| ❑ woman                     | female              |
| ❑ horse                     | animal              |
| ❑ tulip, rose               | flower              |
| ❑ green, red, blue          | colour              |
| ❑ run                       | move                |
| ❑ tiger, leopard, lion      | feline              |
| ❑ trumpet, clarinet, violin | musical instruments |



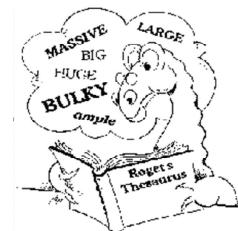
## Metonymy

Substitutes for the object that is meant, the name of an attribute or concept associated with it

- ❑ Kremlin, Whitehall, Washington, Baghdad, Delhi
- ❑ Madison Avenue, Scotland Yard, Downing St., Fleet Street, Wall Street
- ❑ gridiron, ice, turf, diamond, green, pitch
- ❑ "...lend me your ear"

## Synonymy

"Please do not annoy, torment, pester, plague, molest, worry, badger, harry, harass, heckle, persecute, irk, bully, rag, vex, disquiet, goad, beset, bother, tease, nettle, tantalize, or ruffle the animals."



## Synonymy

big, Brobdignagian, colossal, enormous, gargantuan, gigantic, hefty, hulking, humongous, husky, immense, jumbo, Leviathan, looming, lumbering, mammoth, mountainous, ponderous, prodigious, sizable, substantial, tremendous, vast, weighty, whopping

## Synonymy

You should not be aghast, alarmed, amazed, appalled, astonished, bewildered, blown away, bowled over, confounded, dumbfounded, electrified, flabbergasted, flummoxed, overwhelmed, shocked, startled, stunned, stupefied, surprised, taken aback, or thunderstruck at this capacious cornucopia of synonyms in our marvelous English language.

## Synonymy

- automobile car
- marry wed
- fall autumn
- buy purchase

## Partial synonymy

- Deep, profound thought, deep/\*profound water
- Mature, ripe fruit, mature, \*ripe animals

### Synonymous opposites!

- Good scare, bad scare
- Fat chance, slim chance
- Perfectly good = adequate

## Homophony

- bank bank
- date date
- plumb plum
- sea see
- two too
- by bye

|               | Homo-graph                   | Non-Homo-graph |
|---------------|------------------------------|----------------|
| Homo-phone    | Trunk (nose, pants, luggage) | By, buy        |
| Non-Homophone | Lead (be in front, metal)    | moat, lake     |

## Lewis Carrol again

Here the Red Queen began again. 'Can you answer useful questions?' she said. 'How is bread made?' 'I know *that!*' Alice cried eagerly. 'You take some flour--'

'Where do you pick the flower?' the White Queen asked: 'In a garden or in the hedges?'

'Well, it isn't *picked* at all,' Alice explained: 'it's *ground*--'

'How many acres of ground?' said the White Queen. 'You mustn't leave out so many things.'

- Mine is a long and a sad tale!" said the Mouse, turning to Alice and sighing.
- "It is a *long* tail certainly," said Alice, looking down with wonder at the Mouse's tail; "but why do you call it *sad*?" And she kept on puzzling about it while the mouse was speaking, so that her idea of the tale was something like this—

Fury said to a mouse, That he met in the house, 'Let us both go to law: I will prosecute you.' Come, I'll take no denial. We must have a trial: For really the morrow morning I'm nothing if I do not say the mouse to the cook with a trial. With no help, he sat down and said nothing. Till the cook, with a bad

## Polysemy

- is a word or phrase with multiple, related meanings.
- A word is judged to be polysemous if it has two senses of the word whose meanings are related. The vague concept of relatedness is the test for polysemy, judgments of polysemy can be very difficult to make.

- **Mole**
  - a small burrowing mammal
  - consequently, there are several different entities called moles. Although these refer to *different* things, their names derive from 1. :e.g. A Mole burrows for information hoping to go undetected (spy, tumour, pregnancy)
- **Milk**
  - The verb *milk* (e.g. "he's milking it for all he can get") derives from the process of obtaining milk.
- **Wood**
  - a piece of a tree
  - a geographical area with many trees

- **Present**
  - right now, the current moment
  - a gift
  - to show or display (e.g. "SRK was next to present the award")
  - to be physically somewhere (e.g. "Jobs was present at the meeting")
- **Crane**
  - a bird
  - a type of construction equipment

## Antonymy

Words that share all but one semantic property

- **Gradeable:** cold, hot; small, big; happy, sad; fast, slow
  - The words themselves do not provide an absolute scale and are related to the noun they modify; and the negative of one is not synonymous with the other: to not be happy ≠ sad
  - They partition a continuum: minuscule-tiny-small-medium-large-huge-gargantuan; euphoric, elated, happy, so-so, sad, gloomy, despondent
  - More of one is less of the other: wider is less narrow, bigger is less small, taller is less short
  - One of the pair is marked and the other is unmarked: how high (\*low) is the mountain, how tall (\*short) are you, how old (\*young) are you etc.

- **Binary/complementary:** alive, dead; awake, asleep; present, absent

- **Relational:** buy, sell; above, below; give, receive; teacher, pupil
  - Comparative adjectives form relational pairs: if A is taller than B, then B is shorter than A
- **Autoantonyms (contronym or Janus word):** cleave (split, or cling), dust (clear or spread), raise: raze, fast, oversight, shelled, awful, weather

## Two theories

- Componential semantics: meaning broken down into "features"
- Prototypes: "fuzzy" concepts
- *boy*: [+HUMAN, +MALE, -ADULT]
- *man*: [+HUMAN, +MALE, +ADULT]
- *bachelor*: [+HUMAN, +MALE, +ADULT, - WED]
- Meaning as collections of properties/features, typically with two possible values (+/-)

## Features and explanation of lexical relations

- **Anomaly:** *My uncle is female*. feature clash: uncle [+MALE]; female [-MALE]
- **Redundancy:** *That corpse is dead*. shared value for features [-alive]
- **Hyponymy:** *tulip*—flower shared features and extra features
- **Shared meaning:** *boy, girl, colt, filly* all [- ADULT]
- **Antonymy?**

## But

- What is the status of features like +/- EQUINE?
- What is the feature analysis of *hello*, *but*, etc?
- What are the semantic features of e.g. *chair*? *vegetable*?

Sometimes the features are not or cannot be different from the word itself

## Prototypes

- Perhaps at least some words should not be defined in terms of criterial features (features that something must or must not have)
- Some words are best defined in terms of "prototypes"



## Criterion definitions and fuzzy sets

Suppose that senses are defined in terms of features/properties

- Criterion definitions: some/all features are necessary for something to be an X
- Prototype definitions: the more features something has, the more X-like it is (but no one feature is necessary).

## Proper names

- 'My name is Alice, but --'
- 'It's a stupid name enough!' Humpty Dumpty interrupted impatiently. 'What does it mean?'
- 'Must a name mean something?' Alice asked doubtfully.
- 'Of course it must,' Humpty Dumpty said with a short laugh: 'my name means the shape I am -- and a good handsome shape it is, too. With a name like yours, you might be any shape, almost.'

## Proper Names

- Linguistic universal
- Way of describing uniquely to a specific object or entity but has little meaning or sense beyond this power of referral: specific instance of a class
  - Extant names
  - Extinct
  - Fictional
- Proper names are definite, and are typically resistant to being preceded by articles or adjectives and being pluralized

## Phrase and sentence meaning

- 'Then you should say what you mean,' the March Hare went on.
- 'I do,' Alice hastily replied; 'at least--at least I mean what I say--that's the same thing, you know.'
- 'Not the same thing a bit!' said the Hatter.
- 'You might just as well say that "I see what I eat" is the same thing as "I eat what I see"!'
- 'You might just as well say,' added the March Hare, 'that "I like what I get" is the same thing as "I get what I like"!'
- 'You might just as well say,' added the Dormouse, who seemed to be talking in his sleep, 'that "I breathe when I sleep" is the same thing as "I sleep when I breathe"!'
- 'It is the same thing with you,...



## Polysemy in Tamil and Hindi

### □ lagna

|         |             |        |
|---------|-------------|--------|
| answer  | join        | belong |
| cleave  | make        | bulk   |
| feel    | set         |        |
| mind    | settle down |        |
| seem    | sit         |        |
| stick   | sound       |        |
| adhere  | spread      |        |
| accept  | suppose     |        |
| catch   | suspect     |        |
| come by | take        |        |
| come    | tend        |        |
| come to | surmise     |        |
| Be      | cop         |        |
| fall    | connect     |        |
| get     | apply       |        |
| go      | begin       |        |

### kaTTu

1. construct (a house, bridge, etc)/(of birds) build (a nest)
2. inlay (stones in ornaments)/have a denture fixed (in a mouth)
3. compose (a song)
4. build (ship, coach, etc.)
5. heap up earth at the base (of a plant) as support (for a vine)
6. (of a book) make (page)
7. fasten (s.o. with sth); tie (sth.) up; tie (a packet, etc. with a string); bind (sth.)
8. tie (a string or band around part of the body)
9. wear (a sash, diori, etc.); tie (a piece of cloth around the head, waist, etc.)
10. (of a wrist watch, talisman, etc.)
11. bindfold
12. tie a knot (as when getting married)
13. yoke (oxen to a cart, plough, etc.)
14. tie up (flowers); make (a garland)
15. fold hands across the chest or around the knees)
16. many
17. (up water)/stagnate
18. (of phlegm, blood) accumulate; clot
19. (of price, rent, charges) be reasonable
20. pay (a debt, score, etc.)
21. (in a horse race)
22. play (a part in a stage play, film)
23. brand (s.o.); label
24. counteract venom (as an antidote)
25. (of stools) become hard/harden (the stools)

## नामों का सच

- नाम हिलेरी, न तानिक हिले री  
 नाम अगाथा, लिखे गाथा पे गाथा  
 इन सबको देख के घूमा है माथा
- कब, कहाँ कोई क्या कर जाए?  
 नाम के विपरीत रास रच जाए
- नागपुर में नाग बसते नहीं हैं  
 कानपुर में कान पकते नहीं हैं
- पटना में लड़की पटती नहीं हैं  
 विदिशा में दिशा मिलती नहीं हैं

संतरों को खा के संत रोते नहीं हैं  
 मंजु के सर में मन जू नहीं है

बाईट की दुनिया बा-ईट नहीं है  
 मौसीकी से मौसी का रिश्ता नहीं है

चार पाई में मिलती चारपाई नहीं है  
 सैलाना नियासी सैलानी नहीं है

रसगुलों से यारो रस गुल नहीं है  
 और गुलाब-जामुन जो है वो फल-फल नहीं है

## Jokes (with apologies)

- What did the baby pea say to the adult pea?  
 Are you my maTTar?
  - What did the one milk carton say to the other?  
 Whats up doodh?
  - What did one pizza box say to the other?  
 Pizza haT!
  - Teacher: raju, tum kis liye college aate ho?  
 Student: vidya ke khaatir  
 Teacher: toh ab so kyū rahe ho?  
 Student: aaj vidya nahi aayi hai sir
- Homophony lends itself to puns.

## Semantics

### Phrasal Meaning



## Summary of Lexical Semantics

### □ Word Meaning has two components

- Sense (the notion of the meaning of words in the speaker's mind)

Things that don't exist, Santa Claus, Harry Potter or Alice  
 Things that don't have reference: *No woman has become a President of the US; Is there a platypus in your bathroom?*  
*Non-existent, sometimes, either, and, or, the, pity*

- Reference (relationship between the words and the world)

Reference applies not just to nouns but verbs, adjectives etc.  
 Jump > any individual in a set of objects that jump  
 Purple> any individual in the set of objects that are purple

### □ Sense ≠ Reference

The winner of the 2008 US Presidential Election  
 Barack Obama

Pick out the same individual in the world: Image above

1. Obama is married to Michelle
2. The winner of the 2008 US Presidential Election is married to Michelle
3. I want to know if Obama was the winner of the 2008 US Presidential Election
4. I want to know if Obama is Obama

1 and 2 mean the same thing but 3 and 4 don't.

Meaning must be stored in a way that is distinct from reference.

- Dictionary style definitions of word meaning are circular
- Mental image (*mother, Mother Theresa, elephant's mother, bird, the*; prototype theory)
  
- Usage-based definitions – words in context – real-world and grammatical – suitability of word  
Componential semantics (features) and lexical relations (hyponymy, synonymy, antonymy etc.)

## Principle of Compositionality

- The meaning of a phrase or sentence depends both on the meaning of its words and how those words are combined structurally.
- Different meanings with the same words
    - Ram saw Sita
    - Sita saw Ram
  - Different meanings because there are two different structures
    - Visiting relatives can be boring
    - Flying planes can be dangerous

Many lexical relations that we discussed in lexical semantics can also hold between sentences:

- It is easy to play sonatas on this piano PARAPHRASE (synonymous)  
This piano is easy to play sonatas on  
On this piano it is easy to play sonatas  
Sonatas are easy to play on this piano  
To play sonatas on this piano is easy
- I need to buy a pen for Babe (LEXICAL AMBIGUITY, pen, makes S also ambiguous)
- He is alive (He is dead, He is not alive). NEGATION=ANTONYMY

## Noun centred meaning

- (a) adjectives
- Adj N [red balloon]  
property of **redness** added to **balloon** (not otherwise)
  - Det Adj N [the [red balloon]]  
a **particular instance** of a balloon with the property redness
  - Large balloon > large for a balloon
  - Good friend > a kind of friend (modifier properties added to those of the head)
  - False friend > not a friend at all
  - Alleged murderer > unknown whether murderer or not (undetermined)

- (b) PPs: a house with a white fence
- Preposition indicates a relationship between the two objects, the meanings always centre around the meaning of the head
- (c) In compound nouns, the meaning of the noun centres on the head of the compound
- Doghouse vs. housedog (head House vs Dog)
  - Headstrong (Adj)
  - Spoonfeed (Verb)
  - (Head: Right)

## Verb centred meaning

- Thematic roles: the semantic roles played by the arguments (NPs) of a verb and any semantic restrictions placed on them by the verb
- Recall the role of the verb in determining sentence structure – obligatory arguments:  
hit <*one who hits, one who is hit*>  
walk <*a walker*>
- Arguments receive theta roles from the verb; in sentences with movement, connection between the displaced argument and the verb etc. is established by the theta role relationship.

| Thematic role | Description                                       | Example                                       |
|---------------|---------------------------------------------------|-----------------------------------------------|
| Agent         | The one who performs the action                   | Ram wrested the bow                           |
| Theme         | The one or thing on which the action is performed | Ram wrested <b>the bow</b>                    |
| Location      | The place where an action happens                 | It rained <b>in the city</b>                  |
| Goal          | The place to which an action is directed          | Place the books <b>on the shelf</b>           |
| Source        | The place from which an action originates         | He ran <b>from the Institute</b> to his house |
| Instrument    | The means by which an action is performed         | He cut his hair <b>with a knife</b>           |
| Experiencer   | One who perceives something                       | <b>Ravana</b> is angry                        |
| Causative     | A natural force that causes change                | <b>The wind</b> damaged the roof              |
| Possessor     | One who has something                             | The tail <b>of the dog</b> wagged furiously   |

Thematic roles remain unchanged in sentence paraphrases (active-passive). They may also remain the same in non-paraphrases:

- **The boy opened the door with the key**  
AGENT THEME INSTRUMENT
  - **The key opened the door**  
INSTRUMENT THEME
  - **The door opened**  
THEME
  - **The hotel forbids dogs**  
LOCATION
  - **It appears to have rained**  
NO THETA ROLE
- **Theta theory** deals with thematic role ( $\theta$ -role) assignment by (lexical heads) under the ***theta-criterion***.

## The sense and reference of sentences

- Extension (reference), Intension (sense)
- Knowledge of Meaning as Knowledge of Reference and Truth conditions (Sense or Intension)**

What is meaning? The question might be too complex to be tractable. But the semanticist observes that in order to know the meaning of a sentence, one has at least to know under what conditions the sentence is true, i.e. to know its **truth conditions**.

Reducing the question of meaning to the question of truth conditions has proved a very fruitful strategy in studying meaning in contemporary linguistics.

However this cannot be the whole story.

- Consider again the grammatical sentences in the previous slide. Speakers know something more about them - *they know under which conditions each of them is true*.
- And just as we saw in syntax that speakers cannot simply have memorized all the sentences that are grammatical (there is an infinity of them, and speakers are finite creatures), it is certain that they have not memorized their truth conditions either - for the simple reason that there is an infinity of these too.

## Syntactic Knowledge vs. Semantic Knowledge

- Speakers of English can in principle distinguish between an infinite number of grammatical and ungrammatical sentences:
  - a. The rightmost person in the first row is asleep.
  - b. The person immediately to the left of the rightmost person in the first row is asleep
  - c. The person behind the person immediately to the left of the rightmost person in the first row is asleep
  - d. \*Rightmost person in the first row is asleep.
  - e. \*The to the left of the rightmost person in the first row is asleep
- The goal of Syntax is to discover the rules that allow speakers to make such distinctions.

Consider the following series of grammatical sentences:

- a. John could swim
- b. John's father could swim
- c. John's father's father could swim
- d. John's father's father's father could swim

And so on...

- It is clear that no two sentences in this series are **true** under exactly the same conditions.
- And yet given any situation, a speaker of English may (given enough time and patience!) determine whether any given sentence in this list is true.

## An Example of How it all Works

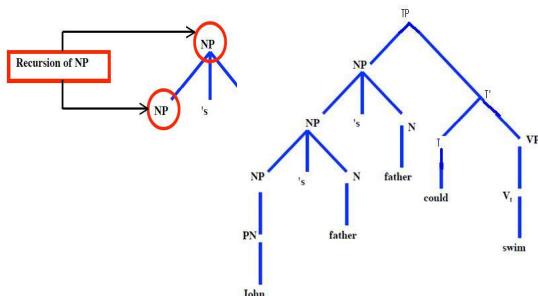
- Consider a toy grammar, one that produces all the sentences in the previous slide:

$$\begin{array}{ll} \text{TP} \rightarrow \text{NP T}' & \text{PN} \rightarrow \text{John} \\ \text{T}' \rightarrow \text{T VP} & \text{N} \rightarrow \text{father} \\ \text{VP} \rightarrow \text{V}_i & \text{T} \rightarrow \text{could} \\ \text{NP} \rightarrow \text{PN}, \text{NP}_1 \text{'s N} & \text{V}_i \rightarrow \text{swim} \end{array}$$

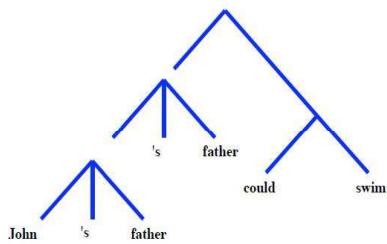
This grammar is similar but not quite identical to those that we discussed in the lectures on phrase structure. The innovation concerns the NP rule  $\text{NP} \rightarrow \text{PN}, \text{NP} \rightarrow \text{NP}' \text{s N}$ .

Here is an instance of recursion of NP. Thus when the 2nd option offered by the rule is applied, we end up with a configuration as in the next slide:

- where recursion of NP is indicated by circles –



- We can recover a simple constituency tree from this phrase structure tree, without labels and non-branching nodes:



- With this background in mind, we can now define some semantic rules that will give truth conditions for each of the sentences below:

- John could swim
- John's father could swim
- John's father's father could swim
- John's father's father's father could swim

- The rules have two parts:

- one set of rules determines what expressions such as *John*, *John's father*, etc. refer to
- the other part determines whether a sentence of the form *x could swim* is true, where *x* is any expression.

## Semantic Rules

The semantic rules in question are:

### (i) Reference

- (i-a) *John* refers to John
- (i-b) For any expression *x*, *x's father* refers to the father of what *x* refers to.



### (ii) Truth

For any expression *x*, *x could swim* is true if and only if what *x* refers to could swim.

(*x*, any individual in a set of objects that swim)



- Suppose that *John's father* was Bill, and that Bill's father was Sam. (*John son of Bill, Bill son of Sam*: Facts about some world)

- According to rule (ii) (**Truth**)  
[[*John's father*] [*could swim*]]

is true just in case what *John's father* refers to could swim.

- But according to rule (i-b) (**NP rule**)

*[John's father]*  
refers to the father of what *John* refers to; and according to rule (i-a) (**NP rule**) *John* refers to John, so *[John's father]* refers to Bill.

Hence in the end the sentence is true just in case Bill could swim – Bill ∈ to the set of objects that swim)

## How about ?



- To compute the reference of the subject *[John's father]’s father* one further step is required. We already established in the preceding example that *[John's father]* refers to Bill.
- Applying rule (i-b) to *[John's father]’s father*, we can now establish that the expression refers to the father of what *[John's father]* refers to, i.e. to Bill’s father who in this story is Sam.
- Hence in the end the sentence *[[[John's father]’s father] [could swim]]* is true just in case Sam Bill ∈ to the set of objects that swim.

- This little procedure, which has only three rules, can be applied to sentences of arbitrary complexity. In fact, it suffices to give the truth conditions of all the sentences below
  - a. John could swim
  - b. John’s father could swim
  - c. John’s father’s father could swim
  - d. John’s father’s father’s father could swim
  - e. John’s father’s father’s father’s father could swim
  - f. John’s father’s father’s father’s father could swimand so on...

- We can see here that syntactic trees have an advantage that was not discussed in previous lectures: they give us a very easy way to define semantic rules (=rules of semantic interpretation).
- In particular, one of the major syntactic constituents, NP, turns out to also be a major semantic constituent, since rule (i-b), which computes the **reference**, is defined for NPs.
- This fact is very general: the constituents created by the syntax are the natural units for the computation of *reference* and *truth conditions*.

- The truth/falsity of a sentence is its **extension** (connection with the real world) and the truth conditions tell you how you may discover its truth/falsity whether or not you actually know it. This is its **intension** (or sense).
- *Mohan believes that India gained independence on January 1, 1965* is **true** if there is some x, x being Mohan who holds such a belief even though the belief itself is false

### □ Obama is asleep

What are the truth conditions?

An individual designated by the word **Obama** must be in the condition designated by the words **is asleep**

### □ An unicorn called Puff bought seven yellow kites

A horsey creature with one horn called by the name [puff] bought seven objects that are yellow and can be made to fly in the wind. The sentence may be false because unicorns don’t exist (or they don’t buy kites etc.) but we know what the ‘**potential**’ conditions need to be.



## Repeat slide

- Sense ≠ Reference  
The winner of the 2008 US Presidential Election  
Barack Obama  
Pick out the same individual in the world: Image above
  - 1. Obama is married to Michelle
  - 2. The winner of the 2008 US Presidential Election is married to Michelle
  - 3. I want to know if Obama was the winner of the 2008 US Presidential Election
  - 4. I want to know if Obama is Obama
- 1 and 2 mean the same thing but 3 and 4 don’t. Meaning must be stored in a way that is distinct from reference.

- Truth conditions also tell us why we can't simply replace one NP with another:

The winner of the 2008 US Presidential elections was Obama

Barrack Obama is Barrack Obama

The winner of the 2008 US Presidential elections is the man married to Michelle Obama.

All have the same **referent** but mean different things because the truth conditions under which each of the NPs may be evaluated are different. And thus they cannot mutually substitute one another and they may be individually false without falsifying the others.

## Relations between sentences

- Important assumption:** linguistic meaning is truth conditional, i.e., to know the meaning of a sentence is to know the circumstances under which the sentence is true. Relations between sentences are now accounted for by the 'truth conditions' that govern each sentence.

**Illustration:** *The police chased the burglar.*

- There is a group of people responsible for law and order ('the police').
- There is a person who is a thief/someone who takes things that are not his ('the burglar')
- The burglar is a member of the set of things being chased.
- 'The police' belong to the set of things that 'chase burglars'

## Paraphrase

- Two sentences are **paraphrases** of each other if the truth conditions are the same (active-passive etc.)
  - It is easy to play sonatas on this piano PARAPHRASE
  - This piano is easy to play sonatas on
  - On this piano it is easy to play sonatas
  - Sonatas are easy to play on this piano
- I bought a new sofa today / I bought a new couch today
- Lecture / deliver a lecture
- See / perceive
- Modality through auxiliary but also phrasally:
  - Can/be able to
  - May/be permitted to etc.

But not all active passive sentences are paraphrases!

Every boy in this room speaks two languages

Two languages are spoken by every boy in this room

## Accounting for paraphrases

- Tom is young
- Tom is an adolescent

### Truth Condition for 1

Existential (x)

x is named Tom

$x \in \{x | x \text{ is young}\}$

### Truth Conditions for 2

Existential (x)

x is named Tom

$x \in \{x | x \text{ is an adolescent}\}$

**Account:** *young* and *adolescent* are lexical synonyms, the truth conditions are the same and so they remain paraphrases of each other.

Passive  $\langle \text{hit} (x,y) \rangle$  where  $x \in \text{hitters}$  and  $y \in \text{those hit}$

## Contradiction

- The truth of one sentence entails the falsity of the other: Sentence P contradicts sentence Q if and only if the truth conditions for P and Q are such that it is impossible for both P and Q to both be true.

My brother is a bachelor. My brother is married.  
This cockroach is alive. This cockroach is dead.  
I live in Trombay. I don't live in Trombay.

## Accounting for Contradiction

- Charles is a bachelor
- Charles is married

### Truth Condition for 1

Existential (x)

x is named Charles

$x \in \{x | x \text{ is a bachelor}\}$

### Truth Conditions for 2

Existential (x)

x is named Charles

$x \in \{x | x \text{ is married}\}$

**Account:** *bachelor* and *married* are lexical antonyms; the first statement and the second are complementary; so it is not possible for some person, here Charles, to be in both sets.

**Sentential negation:** The earth is round  $\rightarrow \neg$  The earth is not round

## Entailment

- A sentence *p* entails a sentence *q* just in case each conceivable situation in which *p* is true is a situation in which *q* is also true (though the converse need not be the case, i.e. it does not have to be the case that every conceivable situation in which *q* is true is a situation in which *p* is also true).
  - I bought a tulip today. → I bought a flower today.
  - I ate a banana today. → I ate a piece of fruit today.
- Sentence *P* entails sentence *Q* iff the truth conditions of *Q* are included in the truth conditions of *P*
  - Prince is a dog. → Prince is an animal.
  - Prince is a dog. ~ → Coby is a cat.

### Explanation

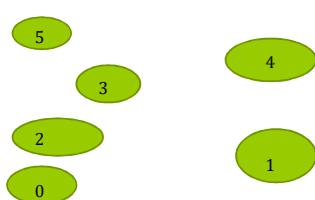
If Prince is a dog, it must also be an animal, because all dogs are animals (hyponym-hypernym). It's possible, however, for Prince to be a dog, and Coby to not be a cat, so no entailment relation.

1. No student came to class ⇒ No student came to class and had a good time
2. No student is over 30 years of age ⇒ No student is over 35 years of age
3. Less than five students came to class ⇒ Less than five students came to class and were happy
4. Less than five students are over 30 years of age ⇒ Less than five students are over 35 years of age
5. Every student came to class ⇏ Every student came to class and had a good time
6. Every student is over 30 years of age ⇏ Every student is over 35 years of age
7. More than five students came to class ⇏ More than five students came to class and were happy
8. More than five students are over 30 years of age ⇏ More than five students are over 35 years of age

9. Every student came to class early ⇒ Every student came to class
10. More than five students are over 35 years of age ⇒ More than five students are over 30 years of age
11. No student came to class early ⇏ No student came to class
12. Less than five students are over 35 years of age ⇏ Less than five students are over 30 years of age

- (i) Whenever *No student VP1* is true, *No student VP2* is true if *VP2* is '**more restrictive than VP1**' (i.e. applies to a subset of the things to which *VP1* applies). For this reason we say that the sentence *No student VP1* is **downward entailing** with respect to its *VP* position.
- (ii) Whenever *Every student VP1* is true, *Every student VP2* is true if *VP2* is '**less restrictive than VP1**' (i.e. applies to a superset of the things to which *VP1* applies). For this reason we say that the sentence *Every student VP1* is **upward-entailing** with respect to its *VP* position.
- (iii) It is then enough to observe that *Less than five students VP1* is also **downward-entailing** in its *VP* position to immediately obtain the patterns seen in the data (1-4). And similarly the pattern in 5-8 can be obtained as soon as it is observed that *More than five students VP1* is **upward-entailing** in its *VP* position

- Less than five students came to class
  - → Less than five students came to class and were happy
- Possible sets of students (less than 5) who came to class:
- Possible sets of students (less than 5) who came to class and were happy



## Accounting for entailments

1. Prince is a dog
2. Prince is an animal

### Truth Condition for 1      Truth Conditions for 2

|                                    |                                        |
|------------------------------------|----------------------------------------|
| Existential (x)                    | Existential (x)                        |
| x is named Prince                  | x is named Prince                      |
| $x \in \{x   x \text{ is a dog}\}$ | $x \in \{x   x \text{ is an animal}\}$ |

Dog is a hyponym of animal (all dogs are animals) so  $\{x | x \text{ is a dog}\} \subseteq \{x | x \text{ is an animal}\}$ ; the truth conditions of 1 are included in the truth conditions of 2; in every case where 1 is true, 2 will also be true. So 1 2

## Structural ambiguity

When more than one meaning results from the way the words and phrases are structured in a sentence, the sentence can be understood to have two or more sets of truth conditions

### Examples

- I read the letter to the editor.
- We raise miniature rabbits and gophers.

## Accounting for ambiguities

□ **Lexical:** A single word in a sentence has different truth conditions simply because the word can be characterized by more than one set. Different truth conditions.

- This is a pen

| Meaning 1                                         | Meaning 2                                 |
|---------------------------------------------------|-------------------------------------------|
| Existential (x)                                   | Existential (x)                           |
| x is something pointed to                         | x is something pointed to                 |
| $x \in \{x   x \text{ is a writing instrument}\}$ | $x \in \{x   x \text{ is a small cage}\}$ |

## Accounting for ambiguities

- **Structural:** modifier scope varies
- **Tiny shells and seaweed are lovely**
- The set *tiny* could intersect with *shells* only or both *shells* and *seaweed*, truth conditions change.

| Meaning 1                                                                                         | Meaning 2                                                                                           |
|---------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| $\{x   x \text{ is tiny}\} \cap \{x   x \text{ is a shell}\} \subset \{x   x \text{ is lovely}\}$ | $\{x   x \text{ is tiny}\} \cap \{x   x \text{ is a shell}\} \subset \{x   x \text{ is lovely}\}$   |
| and                                                                                               | and                                                                                                 |
| $\{x   x \text{ is a seaweed}\} \subset \{x   x \text{ is lovely}\}$                              | $\{x   x \text{ is tiny}\} \cap \{x   x \text{ is a seaweed}\} \subset \{x   x \text{ is lovely}\}$ |

## Presupposition(a special case of entailment)

I'm sorry you arrived late.  
I'm not sorry you arrived late.  
Presupposition: you arrived late.

It was me that broke the lamp  
It wasn't me that broke the lamp  
Presupposition: someone broke the lamp

The mayor of Manchester is a woman.  
P: There is a mayor of Manchester.  
I regret eating your sandwich.  
P: I ate your sandwich.

- A presupposes B: B is part of the assumed background against which A is said (an implicit assumption about the world or background belief relating to an utterance whose truth is taken for granted in discourse).
- A more formal definition of presupposition as a truth relation:
  1. **Step 1:** If **q** (the presupposing sentence) is true, then **p** (the presupposed sentence) is true.
  2. **Step 2:** If **q** is false, then **p** is still true.
  3. **Step 3:** If **p** is true, **q** could be either true or false.

- John's brother has just returned from Texas.  
(P) John has a brother.
- **If q true, then p is true**  
If it is true that John's brother has come back from Texas, it must be true that John has a brother.
- **If q is false, then p is still true**  
If it is false that John's brother has come back from Texas (if he is still there, for example), the presupposition that John has a brother still survives.
- **If p is true then the status of q may be false or true**  
If it is true that John has a brother, it doesn't tell us anything about whether he has come back from Texas or not; we just don't know.

## Difference between entailment and presupposition

- I saw my father today. → I saw someone today.  
If we negate the sentence then the entailment is also canceled.
  - I didn't see my father today. ~→ I saw someone today.  
It might be true, we just don't know. Compare this with the presupposition pair:
  - The mayor of Liverpool is in town today.  
(P) There is a mayor of Liverpool.
  - The mayor of Liverpool isn't in town today.  
(P) There is a mayor of Liverpool.
- Negating the presupposing sentence does not affect the presupposition, whereas negating an entailing sentence destroys the entailment (logical inconsistencies).

## Compositional semantics of Ss

- The relationship between meaning and syntactic structure is referred to as the **Principle of Compositionality**: *The meaning of a whole is a function of the meaning of the parts and of the way they are syntactically combined.*
- Note crucially that there are two components to this:
    - What the parts mean
    - How the parts are combined
  - Word meanings normally combine by regular principles that are dependent on sentence structure – this is seen vividly from the exceptional cases in which they do not – **idioms**. Idioms are cases where a sequence of words has a fixed meaning that is not composed out of the literal meaning of its component words by regular principles. (Second language learning – do not translate).

## For any simple S

- NP VP<sub>i</sub>
- The meaning consists of
  - a) The reference of the NP if determinable (Proper Name) or the set of objects that would qualify as the potential referent
  - b) The VP meaning is the set of individuals who *do* that V or have that property
  - c) If the NP is a member of the denotation of VP, then we can establish the truth value of the S

- The rest of these notes are not included in the final.

## Adjectives

- Sometimes it seems like the specification of truth conditions is trivial:
  - The cat is on the mat.
  - The dog is on the mat.
- Different truth conditions
- But what about more complex cases? Consider:
  - The glass is half full.
  - The glass is half empty.



## Glass

- On the face of it, 'half full' and 'half empty' seem to have the same truth conditions.
- **But:** Consider the following examples:
  - The glass is almost half full. (e.g. 48% full of water)
  - The glass is almost half empty. (e.g. 53% full of water) (more than half full)
- These have different truth conditions
  - Assuming that 'almost' is the same in the two sentences, it must be the case that 'half full' and 'half empty' actually have different meanings
  - If these two phrases were not different in meaning, where else could the difference come from??



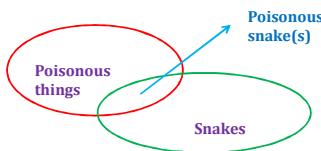
## Other fractions

- As a further point, consider what happens when we replace 'half' by other fractions:
  - The glass is three eighths full.
  - The glass is three eighths empty.
- These do not mean the same thing
- It looks as if 'half full' and 'half empty' mean different things, but sometimes can be true under the same circumstances (optimism vs pessimism)

## More on adjectives

- Some further cases from the study of adjectives illustrate
  - The relevance of our use of sets above
  - The interaction of lexical meaning with compositional meaning
- Let's take another simple example:
  - *poisonous snake*

- **Rule (informal):** When an adjective A modifies a noun N ([A N]), the interpretation of this object is the set defined by the intersection of A's meaning with N's meaning



## Rules

- $\| X \|$  = the set of things denoted by property X
  - **Example:**  $\| \text{red} \|$  = the set of red things
  - This can also be written as  $\{x \mid x \text{ is red}\}$ , read as 'the set of all things x such that x is red'
- How do adjectives and nouns combine by the reasoning above? Because the things that are *poisonous snakes* are the set of things that are (1) poisonous AND (2) snakes.
- $\| \text{poisonous snake} \| = \{x \mid x \text{ is poisonous} \cap x \text{ is a snake}\}$  the intersection of these two sets

The phrase entails both that the object is a snake and that it is poisonous. These are called *intersective adjectives*.

## But...

- Is it always so simple? Consider:
  - Example 1:
    - Larry is a tall guy
    - Larry is a chess player.
    - Therefore: **Larry is a tall chess player** (valid)
  - Example 2:
    - Larry is a skillful artist.
    - Larry is a chess player.
    - Therefore: **Larry is a skillful chess player**. (invalid!)

## So...

- Some other examples:
  - Former bounty-hunter
  - Alleged thief
  - Beautiful dancer
- Do these work in terms of intersection? Are *former bounty-hunters* bounty-hunters at all?

## Consider...

Hilary is a beautiful dancer.

**Meaning1:** She dances beautifully

**Meaning2:** She is beautiful, and she is a dancer  
(she might dance poorly)

**Question:** Do these differences involve different structures, or just a lexically ambiguous set of adjectives?

## Consider also...

Could we have contexts like the following?

- A: There are lots of beautiful dancers here.
- B: Yes, but Mary is the only beautiful beautiful dancer.

If so, which adjective is the one with the 'dances beautifully' interpretation, and which carries the 'is a beautiful person' meaning?

Further:

- Hilary is the only ugly beautiful dancer.
- Hilary is the only beautiful ugly dancer.

**Question (for thought):** Does this mean that the difference is reducible to structure?

These are **subsective adjectives** or relative adjectives and entail only that Hilary is a dancer

## Another type of adjective

John is a former chess player.

- Adjectives like *former* (including *alleged*, *counterfeit*, etc.) are:
  - Not intersective:  
|| former chess player || ≠ ||former|| ∩ ||chess player||
  - Not like *skillful* type adjectives either:  
||skillful chess player|| ⊂ ||chess player||  
||former chess player|| ⊄ ||chess player||
- These are called **intensional adjectives** and map properties to properties <chess player>, <former chess player>
- All of these things are *adjectives*; but how they combine to create larger meanings is determined in part by how they differ from one another
- How to represent such differences goes beyond what we'll do; at this point, we will examine a second factor, syntactic structure

## Intentional adjectives and their properties

### I temporal

- past (former president)
- present (present situation)
- future (future wife)

### II modal

- epistemic (certain victory)
- deontic (necessary ingredient)
- possibility/ability/permission (potential winner)

### III emotive (poor man)

- IV manner (beautiful dancer)
- V object-related (criminal lawyer)

### VI emphaser (outright lie)

## Structure

- One simple case illustrating structural differences involves adjectives combining with nouns
  - (1) in phrases versus
  - (2) in compounds.
- **Example:**
  - Phrase: black board. Meaning: **intersective** (a thing that is both black and a board)
  - Compound: blackboard. Meaning: thing that we write on with chalk. **Not intersective!** A blackboard could be e.g. green.
- How things are put together is crucial.
- This recapitulates what we saw in our study of word structure and syntax

## Structure contd.

□ Here are some further examples:

- John hammered the metal
- John hammered the metal flat

In the second sentence, the adjective *flat* defines the state that the metal moves towards by being hammered.

□ Now, how about:

- John hammered the metal
- John hammered the metal clothed

In the second sentence here, we understand the adjective as defining the state that John was in when he undertook the hammering of the metal

□ However, the structural position of cases like the *clothed* adjective is compatible with either subject and object:

- John talked to Bill clothed (=John or Bill)

- When things like *clothed* appear in the VP, they can be interpreted with either the subject or the object, if it makes sense
- Interestingly:** Further examples show that the *flat* type adjectives and the *clothed* type are in different syntactic positions:
  - John hammered the metal flat clothed
  - \*John hammered the [metal clothed] flat

The second example is deviant because it seems that the first of the two adjectives must go with the object; and in this case, that doesn't make sense.

- When it comes to building meanings, two primary factors must be taken into account:
  - What the individual elements (e.g. specific classes of adjectives in the examples above) mean
  - What the syntactic structures these elements appear in are

## Quantifiers

- Every student read some book.
  - Reading1:** Every student read some book or another (different books)
  - Reading2:** Every student read the same book
- Such ambiguities arise in other cases as well; consider:
  - A student is certain to solve this problem.
    - Reading1:** Some student or the other is going to solve this problem
    - Reading 2:** A particular student is going to solve this problem (e.g. Mary)
- In order to explain the nature of this ambiguity, we will look at some simple logic

## Interpreting Quantifiers

- Understanding the nature of the problem here requires some assumptions about quantifiers.
- In logical analysis, quantifiers are interpreted with respect to some *domain*; think of this as a world. We'll introduce a restricted world below.
- Quantifiers don't seem to refer to things in the way that things like *cat* do. Consider:
  - No students went to the library.
- What would *no students* refer to??

## A restricted domain

- Let's illustrate with respect to a simple domain how the quantifiers work.
- We have a domain (in this example, a set of characters from *Sesame Street*):



## Some basic logic

- In our logic, we need names for individuals. For example:
 

|       |   |          |
|-------|---|----------|
| ernie | ⇒ | Ernie    |
| bb    | ⇒ | Big bird |
| elmo  | ⇒ | Elmo     |
- We also need *predicates*, which are sets of individuals: e.g., *red*, *blue*, *googly-eyed*; these apply to one argument (see below)
- These predicates represent sets, like in our adjective examples; in this cartoon world:
  - $\{\text{blue}\} = \{\text{grover, cookie monster...}\}$
  - $\{\text{googly-eyed}\} = \{\text{cookie monster}\}$

We can then write simple statements, and judge whether or not they are true with respect to our model.

## Example statements

- Some things that we could say (with truth value)
  - Blue(cm) 'cookie monster is blue'; true
  - Red(bb) 'big bird is red'; false
- And so on
- We can also have predicates with two places; e.g.  $Taller(x,y)$  for 'x is taller than y':
  - Taller(bb,cm) true ('big bird is taller than cookie monster')
- How are we going to say things like '*some things* are red', '*no thing* is a chess player', and so on? This is where we need a way of representing quantifiers



## Universal Quantification

- **Universal Quantification ( $\forall$ )**: Represents in logic the meaning of 'every' or 'all'.  
**Example:** (let the predicate Ses be 'is a Sesame Street character')
  - $\forall x \text{ Ses}(x)$
  - This is read as 'for all x, x is a Sesame Street character'
  - This is true in our model, but not in other models, e.g. the real world.
- Meanings like 'no' involve the quantifiers above and negation

## Representing the readings

- **Reading 1:** Everyone saw some person or other
  - $\forall x \exists y (\text{Saw}(x,y))$
  - Read as: 'For all x, there exists some y such that x saw y'
- **Reading 2:** Everyone saw the same person.
  - $\exists y \forall x (\text{Saw}(x,y))$
  - Read as: 'There exists a y such that for all x, x saw y'
- The question for research in natural language semantics is how a single sentence/structure like that of *Everyone saw someone* can have or correspond to these distinct logical representations

## Two quantifiers

- Quantifiers come with *variables*, presented here as x, y, etc.
- **Existential Quantification:  $\exists$**   
It is read as 'there exists an x such that...'  
**Example:**  $\exists x \text{ BLUE}(x)$ 
  - This means 'there exists an x such that x is blue'
  - In our model, this is true
- The other quantifier we need is one that says 'every...'

## Returning to the ambiguous sentence

- Every student read some book
- To simplify, we'll look at:
    - Everyone saw someone  
Which has the same ambiguity
  - The two readings have unambiguous statements



## Utterance meaning ≠ literal meaning

- **Literal meaning:** what is literally said, i.e., the truth conditions of the utterance → job of semantics
- **Implied or contextual meaning:** what is implied from the literal meaning, based on the rules of how conversations work, and the larger context → job of pragmatics

**Example:**

A: I sure could use a cup of coffee  
B: There's a place around the corner called Joe's.  
**Note:** Implied/not literal: Joe's sells coffee

## Intuitions about non-truth conditional meanings

- **Linguistic context matters:** the meaning of the sentence can be 'filled in' by prior utterances
- **Real world knowledge matters:** attitudes and beliefs can further refine meaning
- **Discourse context matters:** the way information is structured in a larger discourse can convey additional assumptions, e.g., new/old information, topics
- **Conversational implicatures:** the rules of co-operative conversations can produce additional meanings, called conversation implicatures

## Linguistic context matters

**Example:** He is there now.

- **Observation:** this example includes several 'deictics', forms that depend on a particular linguistic context (location, time, relation to speaker/hearer).
- **Providing a context:**  
A: I didn't see George earlier at the office  
B: He was out to lunch. He is there now.
- **Contextual meaning:** *he = George, there = the office, now = present time, after lunch*

## Linguistic context vs. real world knowledge

- **Problem:** prior utterance helps fill in some of the information, but not all. Who, among all the Georges of the world, is this George? When is lunch time?
- **Real world knowledge:** the participants in this exchange have certain assumptions about their world that fill in additional meaning, e.g., George is [their colleague, red hair, glasses...], lunch time is around 12:30, etc.

## Examples

- 1. The judge denied the prisoner's request because he was cautious.
- 2. The judge denied the prisoner's request because he was dangerous.

**Task:** determine the most likely meaning of *he* in each case, and identify the real world knowledge that supports this meaning.

## Refinements about truth conditions

1. In isolation: He is there now.
2. Situated in context (linguistic and real world), truth conditions change

A: I didn't see George earlier at the office  
B: He was out to lunch. He is there now.

| Truth Conditions for 1                | Truth Conditions for 2                                      |
|---------------------------------------|-------------------------------------------------------------|
| Existential (x)                       | Existential (x)                                             |
| x is male                             | x is male, called George [colleague, red hair, glasses ...] |
| x is at some location                 | x is at the office [defined by context]                     |
| x is at said location at present time | x is at office at present time [after lunch]                |

## Co-operative Principle

### Maxim of Quality: Truth

- Do not say what you believe to be false.
- Do not say that for which you lack adequate evidence.

### Maxim of Quantity: Information

- Make your contribution as informative as is required for the current purposes of the exchange.
- Do not make your contribution more informative than is required.

### Maxim of Relation: Relevance

- Be relevant.

### Maxim of Manner: Clarity

- Avoid obscurity of expression.
- Avoid ambiguity.
- Be brief ("avoid unnecessary prolixity").
- Be orderly.

## Violation of maxims

Dear Colleague,

Mr. John Smith has asked me to write a letter on his behalf. Mr. Smith is unfailingly polite, is neatly dressed at all times, and is always on time for his classes.

Sincerely yours, ...

- (i) In a letter of recommendation, the professor is normally supposed to mention the most positive features of the student.
- (ii) Homer only mentioned that Smith is polite, neatly dressed and always on time.
- (iii) Therefore these are probably his most positive qualities, and therefore he is probably a bad student.

It is inferred from a reasoning based on (a) what was said, and (b) what a speaker (or here a writer) who tries to communicate as effectively as possible is expected to do given the circumstances.

- Mother: A lot of people are depending on you.
- Daughter: Thanks, that really takes the pressure off.
- **Maxim Violated:** Quality; knowing that "a lot of people are depending on you" does not, in fact, take the pressure off. Meredith is saying something obviously untrue.
- **Implication:** By saying something clearly untrue, Meredith is implying that the *opposite* is true (sarcasm). The true meaning being expressed here is probably more like "That really puts a lot of pressure on me" and perhaps, by extension, "Stop pressuring me."

