Linguistics 100 (Fall 2017) Course Notes

Instructor: Peter Jenks

Notes by Frank Lin

Lecture 1.1: Introduction to Linguistic Science

How did language evolve?

What is language?

Evidence for the sudden emergence of language

- The "Cognitive Revolution" 70,000 BP
- Sudden appearance of complex tools, arts, cultural artifacts
- Ability to hunt large game
- Sudden emergence from Africa

Hypothesis: "Cognitive Revolution" ⇒ language, due to an unexpected mutation in *homo sapiens* that changed brain structure

Language is a biological trait - Eric Lenneberg, *Biological foundations of language* (1967)

- Every normal human being can learn language
- Language learning is automatic in children
- Non-human primates don't have language

Emergence of language transformed primitive *homo sapiens* into modern humans

What is language?

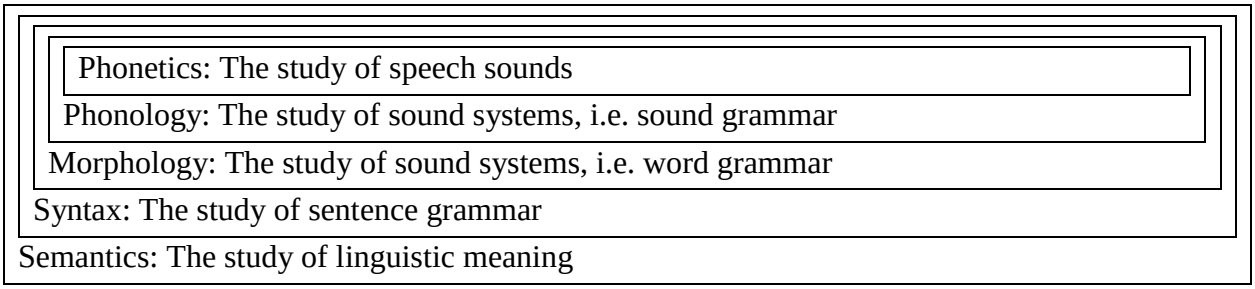
- Language is a biological trait
- Language is unique to *homo sapiens*
- Language emerged suddenly in our evolutionary history

Linguistics examines language as a biological trait, i.e. linguistics is the scientific study of language

Properties of Language

- Language is modality-independent
- Language is generative
- Language is complex (makes use of multiple levels of structure)

Language is complex



Linguistics vs. language prescription

- Linguistics is descriptive: linguists are interested in language as it is used and as it exists in the world

No language is "better" than any other, just as lions aren't "better" than tigers
Language prescription is the practice of telling people how to speak correctly based on some normative standard
The "de-evolution" myth: language is devolving from a pristine earlier state
The "correctness" myth: language is spoken correctly only by a portion of the population
Ex. our speech is just as (if not more) descriptive as the English used in Shakespeare's time

Lecture 1.2: The International Phonetics Alphabet

Written English is a terrible writing system

Design principles for a good spelling system

1. One symbol = one sound (or phone)
2. Separate symbols for separate phones
3. Don't get carried away with predictable differences

One symbol = one phone

1. Many-symbol-to-one-phone problem: fish, thing, battle, syllabus
2. Unpronounced symbols: eight, thistle, lamb, kite
3. Multiple symbols for one phone: sea, see, scene, receive, thief

Separate symbols for separate sounds

1. thing vs. that
2. pin vs. pine
3. who vs. cot
4. one vs. phone

Don't get carried away with predictable differences (English is OK here)

1. keep vs. kool
2. lick vs. lock

Why is English messed up?

The Great Vowel Shift (1350-1700)
English needs spelling reform, but it's not happening... yet

Why writing is unnatural (i.e. not a biological trait)

1. Children have to be taught to read/write, but not to speak
2. Writing doesn't accurately represent linguistic systems
3. Alphabetic writing isn't necessary for written communication
4. Many people have learning disorders that make reading/writing more difficult, e.g. dyslexia

International Phonetic Alphabet

Standard set of symbols for representing speech sounds
Empty boxes: articulatorily possible, yet untested
Shaded boxes: articulatorily impossible

Types of sounds

Categories of speech sounds

Suprasegments vs. segments

Suprasegments: stress and tone

Segments: consonant and vowels

Consonants vs. vowels

Consonants: produced with a constriction in the vocal tract

Vowels: produced without a constriction in the vocal tract

Babbling as a stage in child development and a state in language acquisition

Syllable structure

How many consonants does English have?

25-26, depending on dialect plus glottal stop

Pretty normal size, maybe slightly large

How many vowels does English have?

11-12, depending on dialect

A lot compared to other languages

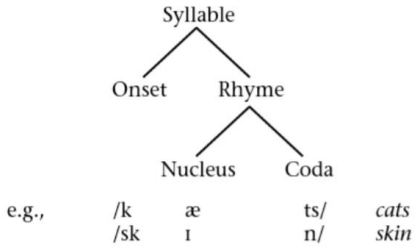
Lecture 2.1: Articulating English Consonants

Consonants vs. vowels

Consonants: produced with a constriction in the vocal tract

Vowels: produced without a constriction in the vocal tract

Consonants vs. vowels in syllables

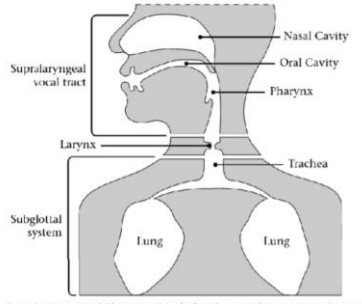


Onset: only consonants

Nucleus: almost always vowels

Coda: only consonants

The vocal tract



From Lieberman and Blumstein, *Speech physiology, speech perception, and acoustic phonetics* (1988), p. 4.
Copyright 1988 Cambridge University Press. All rights reserved. Reprinted with permission.

Normal talking

1. Created by pulmonic egressive airflow (i.e. exhaling)
2. More pressure = more intensity/loudness
3. Airflow can vibrate the larynx, leading to voicing
4. Consonants involve a constriction in the vocal tract, blocking or impeding the airflow

Consonants

CONSONANTS (PULMONIC) © 2005 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

Phonetic transcription

- Represents spoken language
- Differs for different individuals or dialects
- Phonetic transcription should be enclosed in square brackets

[ʃʊd bi ɛnkloʊzd ɪn skwejɹ bɹækɪts]

Three factors in consonant articulation

- Voiced vs. voiceless
- Place of articulation
- Manner of articulation

Voicing caused by the larynx (https://www.youtube.com/watch?v=hfOZxJnY4c8)

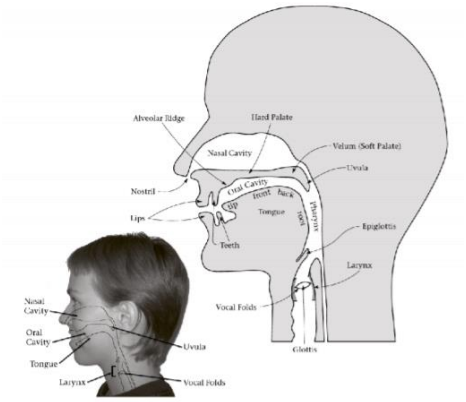
Voiced vs. voiceless consonants

- [f] vs. [v], [θ] vs. [ð], [s] vs. [z], [ʃ] vs. [ʒ]
- [tʃ] vs. [dʒ], [p] vs. [b], [t] vs. [d], [k] vs. [g]

Place of articulation

- Anterior (front) → posterior (back)
- Bilabial, labiodental, dental, alveolar, postalveolar, retroflex, palatal, velar, uvular, pharyngeal, glottal

Articulators



Places of articulation in English

Articulation place	Description	Voiceless ex.	Voiced ex.
Bilabial	Two lips	[p]	[b]

Labio-velar	Two lips + velar constriction		[w]
Labio-dental	One lip and teeth	[f]	[v]
(Inter)dental	Tongue between teeth	[θ]	[ð]
Alveolar	Tongue on alveolar ridge	[t], [s]	[d], [z]
Post-alveolar	Tongue behind alveolar ridge	[ʃ]	[ʒ]
Palatal	Middle of tongue at soft palate		[j]
Velar	Back of tongue at soft palate	[k]	[g]
Glottal	Glottis totally constricted	[ʔ], [h]	

Manner of articulation

Closed ↓ open

Manners of articulation in English

Stop/plosive	No air can pass through mouth	[p, b, t, d, k, g]
Nasal (stop)	Oral stop plus open nasal passage	[m, n, ŋ]
Tap	Very brief oral stop	[ɾ]
Fricative	Turbulent obstruction of air in the mouth	[f, v, θ, ð, s, z, ʃ, ʒ, h]
Affricate	Stop + fricative	[tʃ, dʒ]
Lateral/approximant	Air passes freely on sides of tongues	[l]
Liquid/approximant	Slight constriction without turbulence	[ɹ]
Glide/approximant	Very slight, vowel-like constriction	[w, j]

Most consonants in the English language are unrestricted in their distribution in words

Restricted distribution of [ŋ] - only appears syllable-finally in English (as a coda)

Restricted distribution of [h] - only appears syllable-initially in English (as an onset)

Rarity of [ʒ] - mostly appears in French borrowings or word-medially

Tapping in English - *t* and *d* are often pronounced [ɾ] between vowels in unstressed syllables

Syllabic consonants - in unstressed syllables, [ɹ], [l], and [n] are often syllabic [ɹ̩], [l̩], [n̩]

Glottalization before [n̩] - typically restricted to exclamations (`uh-oh!' [ʔʌʔoʊ]), *t* and *d* are pronounced pronounced [ʔ] before [n̩] in American English dialects *button* [ˈbʌ.ʔn̩], *kitten* [kɪ.ʔn̩]

MRI videos of consonants (https://www.youtube.com/playlist?list=PLKGCiuMS-nxoyh3I9dFG3kJyP4E9zAyXB)

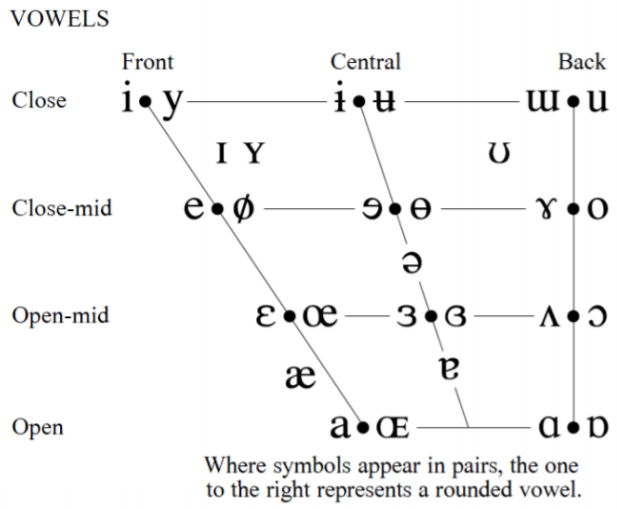
Interactive sagittal section (http://smu-facweb.smu.ca/~s0949176/sammy/)

<u>Lecture 2.2: Articulating English Vowels</u>

Vowel basics

1. Vowels are (almost) always voiced
2. Vowels lack any obstruction to the airstream
3. Vowels are distinguished by the shape of your mouth when pronouncing them

Vowel chart/vowel diagram



English vowels (monophthongs)

[i], [ɪ], [u], [ʊ], [ɛ], [ə], [ʌ], [ɔ], [æ], [ɑ]

Factors in vowel articulation

1. Tongue height
2. Tongue advancement
3. Rounding
4. Tenseness

Tongue advancement

Manner	Description	Examples
Front	Tongue near front of mouth	[i, ɪ, ɛ, eɪ, æ]
Central	Tongue slightly retracted	[ʌ, ə]
Back	Tongue back in mouth	[u, ɔ, ɑ]

Ultrasound of tongue position (https://www.youtube.com/watch?v=b9URTdEe_Ko)

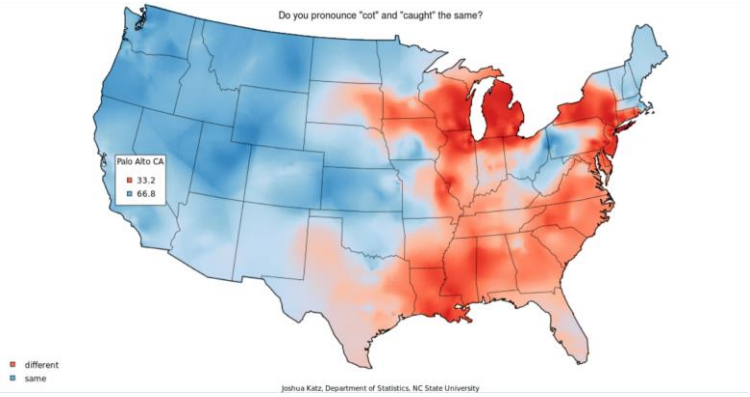
Tongue height

Height	Description	Front	Back
High	Jaw nearly closed, tongue near roof of mouth	seat [sit]	suit [sut] soot [sʊt]
Mid	Jaw slightly open, tongue suspended in mouth	set [sɛt]	saw [sɔ]
Low	Jaw open, tongue low in mouth	sat [sæt]	sot [sɑt]

Many occurrences of [ɔ] are dialectally restricted

[ɔ] has merged with [ɑ] in many American English dialects

cot [cɑt] - caught [cɔt] merger

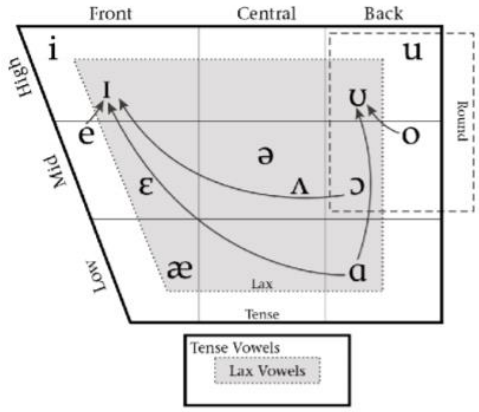


Rounding

- Lip rounding is used for phonetic enhancement in English, i.e. it is redundant
- Rounding always occurs on mid and high back vowels

In other languages, rounding is a contrastive feature of vowels

English diphthongs



[eɪ] - bait [beɪt]
[ɑɪ] - bide [bɑɪd]
([əɪ] - bite [bəɪt])
[ɔɪ] - void [vɔɪd]
[oʊ] - boat [boʊt]
[aʊ] - bout [baʊt]

English diphthongs: open to closed

- [e] only occurs in the diphthong [eɪ]
- [o] only occurs in the diphthong [oʊ]
- Diphthongization seems to be an enhancement of the tenseness of these vowels in English
- [əɪ] is not listed in the chart; it only occurs before voiceless stops

Tense vs. lax

- Tenseness refers to the degree to which a vowel is articulated
- Tense vowels have a more extreme articulation than lax vowels
  - beat [bit] vs. bit [bɪt], suit [sut] vs. soot [sʊt]
  - bait [beɪt] vs. bet [bɛt], coat [koʊt] vs. caw [kɔ]

Cardinal vowel systems (http://www.phonetics.ucla.edu/vowels/chapter3/table3.html)

Five cardinal vowels: [a], [e], [i], [o], [u]

Examples: Spanish, Hawaiian, Swahili, Japanese

[ʌ] ('carat') vs. [ə] ('schwa')

[ʌ] is characteristic of stressed syllables in English

but [bʌt], cut [kʌt]

[ə] is characteristic of unstressed syllables in English

another [ə.ˈnɑ.ðɨ̩] , enough [ə.ˈnʌf]

Suprasegmental features

Features are superimposed on top of segments or syllables

English has stress

	Stressed-unstressed (noun)	Unstressed-stressed (verb)
incline	[ˈɪn.klɑɪn]	[n̩ˈklɑɪn]
intern	[ˈɪn.tɻn]	[n̩.ˈtɻn]
convict	[ˈkɑn.vɪkt]	[kən.ˈvɪkt]
proceeds	[ˈpɹoʊ.sidz]	[pɹə.ˈsidz]
record	[ˈrɛ.kɻd]	[rə.ˈkɔɹd]

Mandarin has tone

Segments	**Tone Numbers[2]**	**Tone Pattern**	**Gloss**
[ma]	55	high level	'mother'
[ma]	35	high rising	'hemp'
[ma]	214	low falling rising	'horse'
[ma]	51	high falling	'scold'

Stress vs. tone

Stress properties	Tone properties
Property of syllables	Property of vowels
Realized as pitch, loudness, length	Realized as pitch only
One main stress per word	Longer words can have multiple tones

Many languages (including Mandarin) have both stress and tone

Lecture 2.3: Sounds of the World's Languages

Non-English sounds that one should know

1. Velar fricatives [x ɣ]
2. Uvular stops [q ɢ] and uvular fricatives [χ ʁ]
3. Alveolar trill [r] and uvular trill [ʀ]
4. High, mid, low front rounded vowels [y, ø, œ]

Some European 'r's

1. French 'r' is a voiced uvular fricative [ʁ] (or trill)
2. German 'r' is a voiced uvular trill [ʀ]
3. Spanish 'rr' is a voiced alveolar trill [r]
4. Spanish 'r' is a voiced alveolar tap [ɾ]

Phonological inventories

1. A language's phonological inventory is the number of contrastive speech sounds that it has
2. For example, [ɾ] is not in the phonological inventory of English because it never contrasts with [t] or [d]
3. Languages vary widely in the size of their phonological inventories

The smallest inventories: Hawai'ian [Austronesian]

Consonants: stops [p, t/k, ʔ], nasals [m, n], approximants [w, l, h]

Vowels: [i, e, a, o, u]

Typically results in longer words

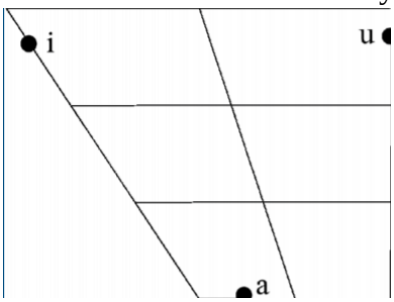
Even smaller: Rotokas [N. Bougainville: Papua New Guinea]

The largest consonant inventories: Adyghe [Caucasian]

		Labial		Alveolar			Post-alveolar		Alveolo-palatal	Retroflex	Velar			Uvular		Pharyngeal	Glottal	
		plain	lab.	plain	lab.	lat.	plain	lab.			plain	lab.	pal.	plain	lab.		plain	lab.
Nasal		m		n														
Plosive	voiceless	p		t							k[1]	kʷ	(kʲ)[2]	q	qʷ		ʔ	ʔʷ
	voiced	b		d							g[1]	gʷ	(gʲ)[2]					
	ejective	pʼ	pʷʼ	tʼ	tʷʼ							kʷʼ	(kʲʼ)[2]					
Affricate	voiceless			t͡s	t͡sʷ		t͡ʃ			ʈ͡ʂ								
	voiced			d͡z	d͡zʷ		d͡ʒ											
	ejective			t͡sʼ			t͡ʃʼ			ʈ͡ʂʼ								
Fricative	voiceless	f		s		ɬ	ʃ	ʃʷ	ɕ	ʂ	x			χ	χʷ	ħ		
	voiced	v[1]		z		ɮ	ʒ	ʒʷ	ʑ	ʐ	ɣ			ʁ	ʁʷ			
	ejective					ɬʼ	ʃʼ	ʃʷʼ										
Approximant									j			w						
Trill				r														

Vertical vowel systems (few vowels due to so many consonants)

Modern Standard Arabic: only three vowels: [i]. [a], [u]



Complexity in phonological inventories

1. Complexity is a property of the entire system
2. Paradigmatic complexity: many distinctions can be made at a single position, or vertically, e.g. onset
3. Syntagmatic complexity: more distinctions can be made simply by adding more structure/content, e.g. longer words

Tricks with airflows: clicks, ejectives, implosives

Non-pulmonic consonants

CONSONANTS (NON-PULMONIC)

Clicks		Voiced implosives		Ejectives	
ʘ	Bilabial	ɓ	Bilabial	’	Examples:
ǀ	Dental	ɗ	Dental/alveolar	pʼ	Bilabial
ǃ	(Post)alveolar	ʄ	Palatal	tʼ	Dental/alveolar
ǂ	Palatoalveolar	ɠ	Velar	kʼ	Velar
ǁ	Alveolar lateral	ʛ	Uvular	sʼ	Alveolar fricative

Ejectives: Montana Salish [Salishan]
Implosives: Igbo [Niger-Congo: Nigeria]
Clicks: Nama [Khoe: Namibiia, Botswana]

Word length and complexity
Swadesh lists - classic compilation of basic concepts for the purposes of historical-comparative linguistics

Lecture 3.1: Phonetics Review, Wrap-up
How many syllables do these words have?

1. *squirrels* - 2
2. *mystery* - 2
3. *burrito* - 3
4. *difference* - 3
5. *laboratory* – 4 (Note: British speakers pronounce ending as *tree*)

Notes on English stress

1. Monosyllabic words: the only syllable is stressed
   [ˈskwɻlz]
2. Disyllabic words: either syllable can be stressed, but tendency for the first syllable in English
   [ˈbɹi.ɾoʊ], [ˈmɪs.tɹi]/[ˈmɪs.tʃɹi]
3. Trisyllabic words: can have one or two stress; if two, will have one main stress (ˈ) and one secondary stress (ˌ)
   [bəˈɹi.ɾoʊ] (one stress), [ˈmɪs.təˌɹi] (two stresses)
4. More than three syllables: will typically have at least two stressed syllables
   [ˈlæ.bɹəˌtɔ.ɹi] (American English), [ləˈbɔ.ɹə.tʃɹi] (British English)

Variation in California Englishes
Vowels
prize vs. price
ride vs. write
flaw vs. flop

SAE vs. California shifted vowels
- bit: [bɪt] vs. [bɛt]
- bet: [bɛt] vs. [bæt]
- bat: [bæt] vs. [bat]

SAE vs. African American Vernacular English
- mind [mɑɪnd] vs. [mɑɪn]
- street [stɹit] vs. [skɹit]
- there [ðeɪɹ] vs. [ðəɹ]/[ðɹ]

Reflections about articulatory phonetics
1. (Spoken) language is speech, not writing
2. Articulation is subject to the physical limits of the vocal tract
3. Different languages use different parts of the vocal tract

Some phonetic universals (for spoken languages)
1. Languages use pulmonic egressive airflow
2. Languages have consonants and vowels
3. Consonants involve place/manner/voicing
4. Vowels involve tongue position
5. Suprasegmental features (tone, stress)

More phonetics
1. Articulatory phonetics: the study of speech articulation
2. Acoustic phonetics: the study of the physical properties of speech sounds
    a. Complex waves
    b. Waveform/spectrogram
    c. Source-and-filter model
3. Auditory phonetics: the study of speech perception
    a. Integration of hearing and vision: the McGurk effect (https://www.youtube.com/watch?v=PWGeUztTkRA)
    b. Categorical perception
4. Neuro-phonetics: the study of the neural control of phonetics

<u>Lecture 3.2: Phonetics Wrap-up</u>

The scientific method
1. Careful observation and description
2. Hypothesis formation: sounds should be described by their articulatory features
3. Hypothesis testing: this system is widely applicable across different languages, so we can conclude this is a good model for spoken language
4. Further applications? Phonology

Signed languages
1. Signed languages have the same basic structure as spoken languages
2. Signed languages are at least as complex as spoken languages
3. Signed languages provide crucial clues to a complete understanding of human language

Lecture 4.1: Phonotactics and Syllable Structure

What is phonology?

Phonology - a grammatical system which regulates the structure and distribution of the sound units of language

Asterisk denotes ungrammatical, or impossible relative to the grammar of a specific language, e.g. *[bnɪf], [skɹɑŋk], *[loʊpf]

What is ungrammaticality?

A linguistic expression is ungrammatical if it is impossible relative to the grammar of a specific language

Two ways to think about ungrammaticality:

1. Unattested, i.e. does not occur in any English words
2. Judged to be impossible based on native speaker intuition

"Knowledge of language"

"Knowledge of language" is knowledge of grammar (Noam Chomsky)

Grammars are systems of rules which generate possible sound sequences, words, and sentences

Linguistic theories are theories of this knowledge, i.e. generative grammars

This knowledge gives rise to native speaker intuitions

Desired properties for a phonological grammar

Which sounds are possible, and in which positions?

Which sounds are predictable and which are not predictable?

What is the internal structure of sounds and words?

Phonotactics and syllable structure

Which sounds are possible, and in which positions?

1. What sounds are in the language to start with?
   = What is the language's phonetic inventory?
2. What sequences of sounds are possible?
   = What are the phonotactic constraints?
3. What sounds can go in which positions?
   = What is the syllable structure?

Phonotactics constraints in English

[bɹ] bring, [gl] glean, [mj] music, [kw] quick, [θɹ] three, [fl] fly, [hj] humor, [sw] sweet
*[knip], *[fgip], *[lpip], *[wlip], *[npip], *[psip], *[ɹfip], *[lɹip]

Phonotactic constraints hold relative to a specific position in a syllable

Phonotactic constraints are part of the grammar of a particular language

Syllable structure

English has a huge number of possible syllables

V	*a*	CV	*no*	CCV	*flew*	CCCV	*spree*
VC	*at*	CVC	*not*	CCVC	*flute*	CCCVC	*spleen*
VCC	*ask*	CVCC	*ramp*	CCVCC	*flutes*	CCCVCC	*strength*
VCCC	*asked*	CVCCC	*ramps*	CCVCCC	*crafts*	CCCVCCC	*strengths*

Most languages are more restrictive
- Hawaiian: CV, V
- Indonesian: CV, V, VC, CVC
- Hebrew: CV, CCV, CCVC, CVC, CVCC

The 'universal CV syllable'
- Every syllable must have a nucleus (e.g. a vowel)
- Onsets are more common than codas, more permissive, and sometimes obligatory (e.g. Hebrew, Thai)
- Codas are more restricted than onsets, and sometimes completely prohibited (e.g. Hawai'ian)

Lecture 4.2: Phonemes and Allophones

Minimal pairs: English vowels
- *beat* [bit] vs. *bit* [bɪt]
- *suit* [sut] vs. *soot* [sʊt]

Contrastive distribution
- Two sounds in the same position can produce words with different meanings
- A pair of words illustrating a contrastive distribution is a minimal pair
- Minimal pairs provide evidence for phonemes

Minimal pairs: Thai stops
- [pʰâː] 'cloth' vs. [pâː] 'aunt' vs. [bâː] 'crazy' vs. [fâː] 'ceiling'
- [pʰ], [p], [b], and [f] have a contrastive distribution in Thai
- [pʰ], [p], [b], and [f] are separate phonemes

Phonemes
- The phoneme is the smallest grammatical unit of a language
- Words consist of a sequence of phonemes
- The distribution of phonemes is unpredictable
- Switching out the phonemes in a word will result in a change in the word's meaning

Allophones of /t/ in American English (most dialects)
- [tʰ] [tʰɑɹ] tar vs. [t] [stɑɹ] star vs. [ɾ] [bæɾ̩] batter vs. [ʔ] [bʌʔn] button vs. [t̚] [bæt̚] bat
- [tʰ], [t], [ɾ], [ʔ], [t̚] have a complementary distribution in American English

Allophones of /k/ in American English (most dialects)
- [kʰ] [kʰɑɹ] car
- [k̚] [skɑɹ] scar, [beɪkɹ] baker, [beɪk̚ɹ̩] bacon
- [k̚] [beɪk̚] bake
- [kʰ], [k], [k̚] have a complementary distribution in American English

Allophones of /p/ in American English (most dialects)
- [pʰ] [pʰɑɹ] par
- [p] [spɑɹ] spar, [pʰeɪpɹ] paper, [hæpn̩] happen
- [p̚] [læp̚] lap
- [pʰ], [p], [p̚] have a complementary distribution in American English

Complementary distribution
- Elements that never occur in the same environment

Allophones

Allophones are phonetically distinct realizations of a phoneme

No minimal pairs distinguish them

Their distribution is predictable

They never occur in the same phonological environment

They are in complementary distribution

Phonemic (contrastive) vs. allophonic (complementary) distributions

	Contrastive	Allophonic
Relation to phonemes	Allophones of separate phonemes	Allophones of the same phoneme
Predictability of distribution	Unpredictably distributed	Predictably distributed
How you can tell	Contrastive distribution; minimal pairs	Complementary distribution

Broad vs. narrow transcription

Phonemic: /-/	Phonetic [-]	Orthographic
/tɑɹ/	[tʰɑɹ]	*tar*
/stɑɹ/	[stɑɹ]	*star*
/bætɻ/	[bæɾɻ]	*batter*
/bʌtn̩/	[bʌʔn̩]	*button*
/bæt/	[bæt̚]	*bat*

Free variation: unpredictable alternation between two allophones

No change in meaning

Sounds are in overlapping distribution

[læp] [læp̚] lap, [bæt] [bæt̚] bat, [beɪk] [beɪk̚] bake

Spanish has a nearly perfect phonemic orthography

Spanish consonants

Phonemic contrast between voiced and voiceless stops

No contrast between voiced and voiceless fricatives

Phonemes and allophones of /b/ in Spanish (most dialects)

[b] [beɾ] ver ‘to see', [gamba] gamba ‘prawn’, [β] [kaβajo] caballo ‘horse’

Allophonic realization of /b d g/ in Spanish (most dialects)

Voiced fricatives are allophones of voiced stops in Spanish

Their distribution is predictable

Word initial	[ganaɾ]	‘to win’	[daɾ]	‘to give’	[beɾ]	‘to see’
Pre-nasal	[teŋgo]	‘have (1sg)’	[andaɾ]	‘to walk’	[gamba]	‘prawn’
Inter-vocalic	[paɣaɾ]	‘to pay’	[deðo]	‘finger’	[kaβajo]	‘horse’

Lecture 4.3: Phonological Rules

Productivity

smattle [smæɾl̩], stilf [stɪlf], tarb [tʰɑɹb]

A rule is productive if it can be applied in novel environments

Productivity appears during language acquisition

Productivity in real examples

/t/ → [t̚]	/t/ → [ɾ]
bat [bæt̚]	*batter* [bæɾɻ]
loot [lut̚]	*looted* [luɾəd]

write [rəi̚t]	*writer* [ɹəɪɾɻ]

Phonological rules: the description - /X/ → [Y]

X is the underlying form (phoneme) and Y is the surface form (allophone)

Model: phonemic form ⇒ rules ⇒ phonetic form

Phonological rules: the environment - /X/ → [Y] / environment

"X becomes Y in [environment]"

Adding the environment

/t/ → [tʰ] / word initially	[tʰɑɪ] *tar*
/t/ → [t] / after s	[stɑɪ] *star*
/t/ → [ɾ] / after vowel as onset of unstressed ɹ or vowel	[bæɾɻ] *batter*
/t/ → [ʔ] / after vowel as onset of unstressed syllabic n	[bʌʔn̩] *button*
/t/ → [t̚] / word finally	[bæt̚] *bat*

Special symbols for environments

/ _ X	before X	#	word boundary
/ X _	after X	.	syllable boundary
/ X _ Y	after X and before Y	C	consonant
/ _{X, Y}	before X or Y	V	vowel

Environments using special symbols

/t/ → [tʰ] / #_	[tʰɑɪ] *tar*
/t/ → [t] / s_	[stɑɪ] *star*
/t/ → [ɾ] / 'V._{V, ɻ}	[bæɾɻ] *batter*
/t/ → [ʔ] / 'V._ n̩	[bʌʔn̩] *button*
/t/ → [t̚] / _#	[bæt̚] *bat*

Elsewhere rules

Final version of phonological rules for /t/ in American English

/t/ → [tʰ] / elsewhere	*tar, tricky, attention, entry*
/t/ → [t] / s_	[stɑɪ] *star*
/t/ → [ɾ] / 'V._{V, ɻ}	[bæɾɻ] *batter*
/t/ → [ʔ] / 'V._ n̩	[bʌʔn̩] *button*
/t/ → [t̚] / _#	[bæt̚] *bat*

Natural classes

Natural classes are sets of phonemes with a shared phonological property, such as manner or place of articulation

Phonological rules provide evidence for natural classes as important components of our mental grammar

The most basic natural classes are consonant and vowel

Any combination of articulatory features can be a natural class

Natural classes provide evidence that articulatory features are part of the phonological system

Using natural classes allows us to capture generalizations over classes of similar rules

Some common natural classes

Consonants, vowels

Obstruents - stops, fricatives, (nasals)

Sonorants – vowels, liquids, glides, (nasals)
Labial, alveolar, velar
Sibilants (/s, z, ʃ, ʒ/)

Phonological rule for Spanish voiced stops
/voiced stop/ → [voiced fricative] / V_V
/voiced stop/ → [voiced stop] / elsewhere

Phonological rule for Thai stops and fricatives
/labial obstruent/ → [unreleased voiceless bilabial stop] /_#
/labial obstruent/ → [labial obstruent] / elsewhere

Lecture 5.1: Solving Phonology Problems

Recap on phonemes, allophones, rules
Phoneme - a speech sound which is capable of making meaningfully distinct words in a given language
Allophones - the phonetic realizations of phonemes in distinct phonological environments
Phonological rules - statements which generate allophones from phonemes in particular environments

Recap on natural classes
Natural class - A set of phonemes with a shared phonological property, such as manner or place of articulation
Feature set - The set of phonological features or properties that describe a natural class

Phonemic analysis
Language speakers produce allophones/surface forms
Phonemes cannot be detected with direct empirical observation
Phonemic analysis is a procedure by which phonemes can be discovered from a surface distribution of allophones

Steps to phonemic analysis
1. Determine what sounds are being analyzed
2. Make a list of environments
3. Identify natural classes in the environments
4. Look for complementary gaps in the environments
5. State a generalization about the distribution of each sound
6. Determine the identity of the phoneme and its allophones based on predictability
7. Write the rules

Lecture 5.2: Types of Phonological Rules

An unattested rule: /θ/ → [ŋ] /_i

A theory of rules?
1. What is a possible phonological rule?
2. What is an impossible phonological rule?
3. Where do phonological rules come from?

Phonetic grounding
Most phonological rules are phonetically motivated

Articulatory motivation: easier to pronounce words after rule

Auditory motivation: a phonemic contrast in some environment is undetectable

Types of phonological rules

1.	Assimilation	/X/ → [Y] / _Y	Similarity-based rule
2.	Dissimilation	/X/ → [Y] / _X	Similarity-based rule
3.	Insertion	Ø → [X] / ...	Non-allophonic rule
4.	Deletion	/X/ → Ø / ...	Non-allophonic rule
5.	Metathesis	/XY/ → [YX] / ...	Non-allophonic rule
6.	Strengthening	/Weak/ → [Strong] / ...	Strength-based rule
7.	Weakening	/Strong/ → [Weak] / ...	Strength-based rule

Assimilation: /X/ → [Y] / _Y

Coarticulation - it is easier to pronounce adjacent sounds if they have overlapping articulatory features

e.g. Nasal place assimilation in English

impress	lump	*lunp	/n/ → [m] / _bilabial stop
interrupt	vent	*vemt	/n/ → [n] / _alveolar stop
incredible	blink	*blinp	/n/ → [ŋ] / _velar stop

e.g. Vowel harmony in Turkish

/balta-lar/ [balta-lar] ‘axes’

/palto-lar/ [palto-lar] ‘overcoats’

/ev-lar/ [ev-ler] ‘houses’

/kedi-lar/ [kedi-ler] ‘cats’

/gøz-lar/ [gøz-ler] ‘eyes’

Phonological rule: /a/ → [e] / front vowel C(C)_

Not triggered by immediately adjacent phoneme; such rules are uncommon

Dissimilation: /X/ → [Y] / _X

Avoidance - it is hard to use the same articulation twice in the same word

Dissimilation is less common than assimilation

Dissimilation often applies to non-adjacent phonemes

e.g. /l/~/r/ dissimilation in Georgian (Kartvelian: Georgia)

[p’olon-uri] ‘Polish’

[somχ-uri] ‘Armenian’

[sur-uli] ‘Assyrian’

[p’rusi-uli] ‘Prussian

Phonological rule: /r/ → [l] / preceded by /r/ in the word

Non-allophonic rules: do not involve a relationship between phonemes and allophones

Insertion: Ø → [X] / ...

e.g. voiceless stop insertion in English

hamster	/hæmstɻ/	[hæmpstɻ]
dance	/dæns/	[dænts]
strength	/stɹɛŋθ/	[stɹɛŋkθ]

Phonological rule: Ø → [voiceless stop] / nasal_voiceless fricative

Deletion: /X/ → Ø / ...

e.g. initial /h/ deletion in English

/ɪz hi tɑkɪŋ tu hɛr/?

[izitʰɑkɪntʰuɛr]?

Phonological rule: /h/ → Ø / #_ (in unstressed syllables)

Metathesis: /XY/ → [YX] / ...

e.g. metathesis in Leti (Austronesian: Indonesia)

/danat + kviali/ [dantakviali] ‘millipede'

/ukar + ppalu/ [ukrappalu] ‘index finger’

/ukar + lavan/ [ukarlavan] ‘thumb’

Phonological rule: VC → CV / _CC

Strength-based rules

Sonority - a measure of openness for segments

Sonority scales: (strong) vowel > glide > liquid > nasal > fricative > stop (weak)

English sonority scale

[a] > [e o] > [i u j w] > [ɾ] > [l] > [m n ŋ] > [z v ð] > [f θ s] > [b d g] > [p t k]

Vowels have weakest consonantality

Stops have strongest consonantality

Strengthening: /Weak/ → [Strong] / ...

e.g. Thai final constant neutralization

/pʰ/ → [p] / _#

/p/ → [p] / _#

/b/ → [p] / _#

/f/ → [p] / _#

Phonological rule: /labial obstruent/ → [unreleased voiceless bilabial stop] /_#

Weakening: /Strong/ → [Weak] / ...

Weakening is common between vowels, which are the most sonorous sounds

Weakening is also more common in unstressed syllables

e.g. Spanish voiced stop fricativization

/g/ → [ɣ] / V_V

/d/ → [ð] / V_V

/b/ → [β] / V_V

Phonological rule: /voiced stop/ → [voiced fricative] / V_V

e.g. English flapping

/t/ → [ɾ] / 'V._{V, ɹ̩} [bæɾɹ̩] *batter*

/t/ → [ʔ] / 'V._n̩ [bʌʔn̩] *button*

One final rule: enhancement

Enhancement

e.g. English vowel lengthening

bat /bæt/ [bæ̰t̚]

bad /bæd/ [bæːd]

leak /lik/ [lḭk̚]

league /lig/ [liːg]

leaf /lif/ [lif]

leave /liv/ [liːv]
Phonological rule: /V/ → [Vː] / _voiced consonant
Vowel lengthening in English enhances the word-final contrast in voicing that is otherwise difficult to hear

Lecture 5.3: Historical Linguistics

Introduction to historical linguistics
- Synchronic linguistics - the study of languages at a particular point in time
- Diachronic linguistics - the study of languages over time

Synchronic linguistics began in the late 19th century
Diachronic linguistics began in the late 18th century
Language change acknowledged by Sir William Jones (1786)
Basic principles of historical linguistics
1. Languages are constantly changing
2. Related languages come from a common proto-language
3. Language change is regular, or systematic
4. Language change is gradual, reflected in dialectal variation

Indo-European languages
In historical linguistics, an asterisk is used before reconstructed forms
Sound change

Proto-Indo European (4500-3500 BCE)	*dwoh
Proto-Germanic (750-0 BCE)	*twai
Old English (500-1000 CE)	[twa]
Middle English (~1400 CE)	[twoː]
Modern English	[tʰu]

Phonetic vs. phonological change
- Phonetic change - the change in the pronunciation of an individual phoneme
  - Early Modern English (~1600) [r] > Modern American English [ɹ]
- Phonological change - a change which results in a new phonological system or phoneme
  - Old English (~900) *leaf-es* > Middle English (~1200) *leav-es* / V_V
  - Middle French (~1200) *sauver* > Middle English [savə] ‘save’

Regularity of sound change
- Old English to Modern English
- [luːs] > [lɑʊs] ‘louse’
- [huːs] > [hɑʊs] ‘house’
- [muːs] > [mɑʊs] ‘mouse’
- [uːt] > [ɑʊt] ‘out’
- Sound changes eventually apply to all instances of a sound within a particular language

Conditioned vs. unconditioned change
- Conditioned change - a sound change that only occurs in a particular environment
  - Old English [f] > Middle English [v] / V_V
  - Closely resemble phonological rules
  - Typically have clear phonetic motivation

Phonological rules provide evidence for sound change in action

Unconditioned change - a sound change that occurs to all instances of a particular sound

Old English *u: > Modern English [ɑʊ]

Unlike phonological rules, because they have no environment

Arise from phonetic shifts in pronunciation, usually due to existing variation among speakers

Two examples from history of English: Grimm's Law and The Great Vowel Shift

Grimm's Law

Proto Indo-European > Proto Germanic

PIE *b^h/*d^h/*g^h/*g^{wh} > PG *b/*d/*g/*g^w

PIE *b/*d/*g/*g^w > PG *p/*t/*k/*k^w

PIE *p/*t/*k/*k^w > PG *ɸ/*θ/*x/*x^w

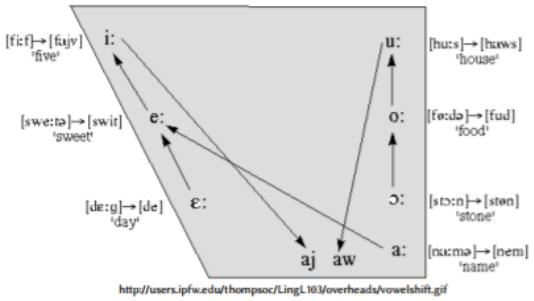
This change is called a "chain-shift"

e.g. Germanic *father* vs. Romance *padre*

Establishes Germanic languages as a group within Indo-European

The Great Vowel Shift

Middle English > Early Modern English



The vowel shift was accompanied by word final schwa deletion

ME *ə > EME Ø / _#

The comparative method

The comparative method has been used to establish genetic relatedness between languages, reconstruct proto-languages, and identify sound-changes

The comparative method is only possible because sound change is regular

Three steps of the comparative method:

1. Collect a cognate set from languages you think are related
2. Determine the correspondences in the language
3. Determine the proto-sound for each position

Computational reconstruction

<u>Lecture 6.1: 'Language and linguistics on trial' (Part 1)</u>

Review of language change and variation

Language change is inevitable and ongoing

Evidence of language change: phonological rules and dialectal variation

Every dialect/language has the same basic structural components

From a historical perspective, the relationship between dialect is "flat"

- Dialects and social class
  - Modern society is stratified into groups with different amounts of power (notably "the ability to effect change" – Martin Luther King, Jr.)
  - People in higher social strata sometimes believe that they are in that position because it reflects a "natural" order (vs. luck, history, oppression)
  - This is reflected in people's attitudes about language: dialects associated with privileged groups are assumed to be inherently superior to dialects associated with poor or marginalized groups
  - Consequently, it is often assumed that people who speak such dialects do so because they are not capable of speaking more prestigious dialects
- Linguistic prescriptivism
  - Linguistic prescriptivism is telling people how they should talk/write
  - Fundamentally, this kind of language reveals classist assumptions about linguistic superiority
  - These assumptions are reinforced as part of many students' education
  - These assumptions also play an important role in society: people make assumptions about other people's intelligence based on how they speak
- Vernacular language in the courtroom
  - 'Language and linguistics on trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond'
  - Two groups of speakers at a disadvantage in legal proceedings
    - Second language or foreign language speakers
    - "Second dialect" speakers (Eades 2010)
  - Eades on legal proceedings in Australia
    - "On most occasions on which Aboriginal speakers of varieties of English give evidence, there is no interpreter and thus no mechanism for dialectal differences to be drawn to the attention of the court."
  - Types of linguistics differences which lead to misinterpretation in Aboriginal English
    - Phonological: "properly his father" transcribed as "probably his father" due to /p/~/b/ variability
    - Lexical/semantic: "half moon" meant "crescent moon", a crucial distinction in a court case about events at night
    - Pragmatic: silence interpreted as uncooperativeness or dishonest, when it is seen as normal to begin a response with silence in Aboriginal English
  - Similar issue with Jamaican Creole in a case in the UK
- African American Vernacular English (AAVE)
  - The language of the African diaspora in the United States
  - Not a single dialect; many distinct variations of AAVE exist across the United States and even in the same regions, though AAVE dialects do have some similarities
  - Shares a number of traits with non-AAVE dialects, e.g. Southern English
- Attitudes about AAVE
  - Often derided in popular media
  - 1997 Oakland School Board endorsed the use of ebonics in the classroom

Triggered an extremely critical response from many figures, e.g. Jesse Jackson

AAVE speakers and the law

African American men are ~six times more likely to be incarcerated than white men in the US

Many issues facing African American defendants in US courts

Rickford and King argue that one is both attitudes about AAVE and courts' lack of proficiency in AAVE

AAVE in the courtroom

In one 1965 case, the defendant spoke AAVE and pleaded self-defense in a murder trial

He was eventually convicted of murder, despite communication issues

Prosecutor: "it would have been easier to bring out the facts of the case if witnesses were unable to speak English, so that competent interpreters could have been used" (ibid.)

Jurors later said that "the greater part of the testimony had been incomprehensible to them"

Transcription issues

Trayvon Martin case and Rachel Jeantel

Rachel Jeantel: key witness for the prosecution

On the phone with Trayvon Martin immediately before his death

Identified a voice which was crying for help as Trayvon Martin's and described Martin's fear of the man following him

Central points of Rickford & King's paper

Rachel Jeantel speaks a relatively typical variety of AAVE

This led people, including the jury, to misunderstand discredit her testimony

Rachel Jeantel as a typical AAVE speaker

Syntactic features

Inflectional/morphological features

3rd singular pres. -s absence

Possessive -'s absence

Plural -s absence

is/*are* absence rates

Phonological features

pin/*pen* merger: [ɛ] > [ɪ]/ _n

Cluster reduction: /fæs/ 'fast'

<u>Lecture 6.2: 'Language and linguistics on trial' (Part 2)</u>

Rachel Jeantel (witness) was a fluent speaker of AAVE

Nevertheless, the jury could not understand Jeantel and discredited her testimony

Rachel Jeantel's linguistic background

Trilingual (fluent in Spanish, Haitian Creole, AAVE)

Influences of Caribbean English and Haitian Creole on Jeantel's AAVE

Response to Jeantel's language was generally critical

Possible factors

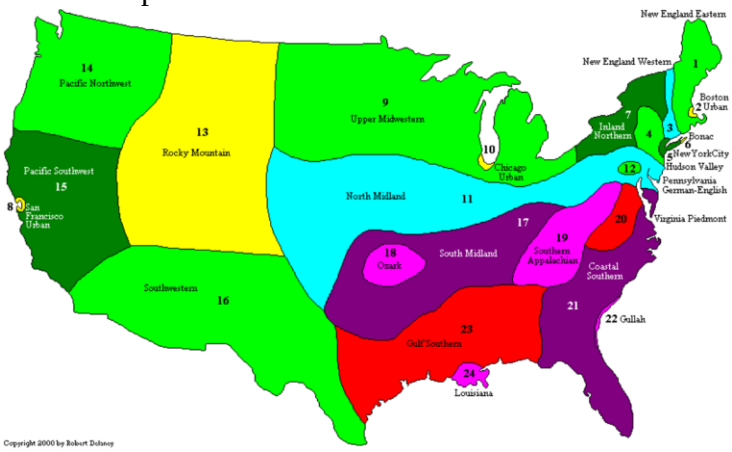
White jury's ability to understand AAVE

Non-native speakers of AAVE often misinterpret AAVE
Jeantel's voice and courtroom acoustics also potentially a factor
Dialect prejudice
Makeup of jury was 5 European-American women and 1 Afro-Dominican woman
Non-native speakers of AAVE often misinterpret AAVE
Post-creole continuum
Acrolect - prestige dialect or language, typically associated with a linguistic "standard"
Basilect - stigmatized dialect or language, typically a vernacular associated with a socially marginalized group
Mesolect - dialects or languages midway through acrolects and basilects
Basilectal and mesolectal native speakers often understand and even speak the acrolect
Acrolectal native speakers rarely fully control lower prestige dialects
Code-switching - the phenomenon of switching between dialects in different social situations
Dialect prejudice
Wolf (1959)
Nembe minority in Nigeria speak and understand prestige Kalabari
Kalabari speakers cannot speak or understand Nembe
Rubin (1992)
Played audio to undergraduates while showing them one of two pictures
When looking at a picture of an ethnic European and ethnic Asian, Asian speech judged as "much more accented" and "much harder to understand" despite the audio being identical
Anti-vernacular bias
Stereotypes against vernacular dialects and local languages are widespread around the world
These attitudes are often strongest among speakers of the dialects and languages
Linguistic competence is fixed early in life (~before 13) and is automatic, so individuals in vernacular-speaking communities are always fluent vernacular speakers
Vernacular English in educational contexts
Written English (Standard English) is based on mainstream British and American English pronunciations (especially syntax/morphology)
This puts speakers of vernacular English at a fundamental disadvantage when learning to read
Dialect-based reading materials have been developed to bridge these gaps, but they are almost never adopted in educational contexts despite clear experimental evidence showing that they effectively bridge the gap

<u>Lecture 6.3: Sociolinguistics</u>
Languages and dialects
Linguistic variety - any form of language characterized by systematic features
Idiolect - the language variety of an individual speaker
Dialect - a form of language spoken by a group of speakers (a linguistic community) which is systematically different from other forms of the same language

Mutual intelligibility - the ability of two speakers from different dialects of the same language to understand what the other is saying

Dialect map of the US



Variation

Linguistic variable - a specific linguistic feature that varies between speakers of different dialects

Isogloss - a boundary that marks the distribution of two variables

Regional dialect variation - variation that is geographically defined

Social dialect variation - variation that is defined based on social factors, including age, gender identity, class, ethnicity, social orientation, etc.

Accent - different phonological features

Two ways to identify dialects

Surveys (http://www.nytimes.com/interactive/2013/12/20/sunday-review/dialect-quiz-map.html)

Sociolinguistic interviews or readings (http://accent.gmu.edu/)

Standards and prestige

Standard dialect - an idealized dialect associated with speakers associated with prestige (privilege or powerful social status)

Nonstandard dialect - a dialect associated with speakers outside the standard

These are social effects; linguistically speaking, no one dialect or language is better, more correct, more systematic, or more logical than any other

e.g. Arabic dialect groups

Arabic languages/dialects have 290 million speakers

Modern Standard Arabic, based on Classical Arabic, is the language of educated Arabic speakers

Dialects and social class

Modern society stratified into groups with different amounts of power

People in higher social strata sometimes believe that they are in that power because it reflects a “natural” order (vs. luck, history, oppression)

This is reflected in people's attitudes about language: dialects associated with privileged groups are assumed to be inherently superior to dialects associated with poor or marginalized groups, and their speakers are "superior" language users

If anything, the opposite is true: speakers of nonstandard dialects are fluent in multiple dialects while native speakers of standard dialects can only understand that dialect

Linguistic variables in society

William Labov - father of sociolinguistics

Identified the role of social factors in variation

Pioneering work on grammatical features of African American Vernacular English

Two famous studies

New York Department Stores

'The social stratification of English in New York City' (1966)

Use of [ɹ] increased in repeated speech, evidence of a formal register

Salespeople at nicer stores had increased rates of [ɹ] use in repetition

Martha's Vineyard

'The social motivation of sound change' (1962)

Demonstrates the role of covert prestige in a rural community, an island in Massachusetts called Martha's Vineyard

Studied the centralization of [ɑw]/[ɑj] to [əw]/[əj]

Used detailed interviews and questionnaires

Martha's Vineyard was (at the time) the poorest county in Massachusetts, heavily dependent on summer tourists

Centralization increased with economic hardship across all ethnicities

Therefore, centralization was a mark of covert prestige

Desire to sound different drove a sound change on Martha's Vineyard

Types of linguistic prestige

Overt prestige - choosing to identify with the standard

Covert prestige - choosing to differ from the standard

Speech register - formal and informal registers

Language and identity

The social meanings of variables

Linguistic variables are resources that speakers use

Speakers actively negotiate their identity using linguistic variables

Speakers adopt different linguistic variables in different environments to take on different "personas"

e.g. The California Vowel Shift

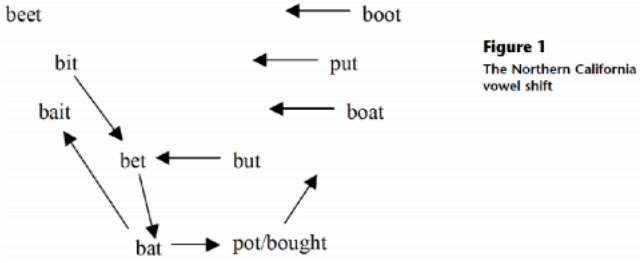


Figure 1
The Northern California vowel shift

Why is the California Vowel Shift happening?
- Language change in action in California
- A marker of a regional "accent"
- (White?) Californians constructing a distinct Californian identity

Dialects and stereotype
- "Features of the California white Anglo dialect have been popularized, gendered, raced, and classed through their media association with the male surfer and the Valley Girl—gendered icons of privilege, materialism, and empty-headedness, but also national trend-setters and the embodiment of white California." Eckert 2008

'Where do ethnolects stop?' (Eckert 2008)
- Examined the distribution of the California Vowel Shift in fifth graders at two elementary schools
- School 1 - ethnically diverse, including a large Chicano population
- School 2 - mostly Anglo
- Observation of the paper: both White and Chicano speech are "ethnolects"
- An ethnolectal split in [æ]
  - The "Standard" California Vowel Split:
    - [æ] → [eə] / _nasal stops
    - [æ] → [a] / elsewhere
  - Chicano English
    - An "ethnolect" associated with the Mexican American community
    - A distinct English dialect (vs. "Spanglish", or code-switching)
    - California Chicano English often has shifted vowels, but:
    - [æ] → [a] (unrestricted)
- White identity is socially negotiated like Chicano or Black identity; standard dialects are ethnolects
- The elementary school students in Eckert's study use [æ] to index or point to a particular social stance
- This process illustrates the social motivation for sound changes

<u>Lecture 7.2: Morphology</u>

Morphology - the study of morphemes and word structure

Morphemes
- A morpheme is the smallest meaningful unit of language
- Complex words consist of a sequence of morphemes
- The same morpheme can appear in different words
- Switching out the morphemes in a word will result in a change in the world's meaning
  - i.e. they have a contrastive distribution

Types of words
- Morphologically simple words - words consisting of a single morpheme
- Morphologically complex words - words consisting of multiple morphemes

Types of morphemes
- Free morphemes - morphemes that function as simple words

Bound morphemes (affixes) - morphemes that must combine with other morphemes to make words

Types of affixes

Prefix - a bound morpheme that precedes a free morpheme

Suffix - a bound morpheme that follows a free morpheme

Lexical category

The features of a word that determines its basic morphological and syntactic distribution

Noun - I like dogs/cats/air/milk/thinking.

Verb - I like/see/eat/admire/examine/treasure dogs.

Adjective - The big/furry/red/dirty/smart dog.

Adverb - I really/almost/barely/quickly eat bread.

Open and closed categories

Open lexical categories - Categories that we can add items to, that we can coin new words for

Closed lexical categories - Categories which are primarily grammatically defined, and which we can't come up with new members for

Pronouns - I/you/he/she/they/we

Determiners - the/a/every/this/that

Conjunctions - and/or/but/slash

Inflectional vs. derivational morphemes

Derivational morphology

Morphology that:

1. changes the lexical category of a word
2. produces a substantial shift in meaning
3. both

V-to-V, Adj-to-Adj:	un-	un-happy, un-tie, un-reliable
V-to-V:	re-	re-tie, re-do, re-verberate

Inflectional morphology

Morphology that:

1. is grammatically controlled
2. has only a minimal semantic effect on the stem

Complete list of English inflectional morphemes (note all suffixes)

Noun inflection:	plural -s	dog-s, cat-s, rose-s
Adjective inflection:	comparative -er	slow-er, fast-er, ugli-er
	superlative -est	slow-est, fast-est, ugli-est
Verb inflection:	progressive -ing	walk-ing, rhym-ing
	past participle -en	eat-en, giv-en
	past tense -ed	walk-ed, rhym-ed
	present tense -s	walk-s, rhyme-s

Note: the possessive -'s is sometimes given as an inflectional affix in English, but since it attaches to noun phrases and not nouns, the result is a different syntactic category, so it is not included here

Derivation + inflection

un-tie-d derivational prefix – stem – inflectional suffix
dirt-i-er stem – derivational suffix – inflectional suffix

In English, words have only one inflectional affix at a time

Steps for morphological analysis:

1. Propose morphological boundaries in complex words
2. Check for each morpheme in other words, revise
3. Identify the meaning which goes with the morpheme
4. Write own the morphemes and their meanings

<u>Lecture 7.3: Allomorphy</u>

Morphemes and allomorphs

Comparative -er

trim [tʰɹɪm] *trimmer* [tʰɹɪm-ɻ]
slick [slɪk] *slicker* [slɪk-ɻ]
white [wəɪt] *whiter* [wəɪɾ-ɻ]
red [ɹɛd] *redder* [ɹɛɾ-ɻ]

The presence of the comparative suffix /-ɹ/ triggers the flapping rule in American English

Phonological rules in the presence of morphology

American English flapping

/t/ → [ɾ] / 'V._{V, ɻ}

Derivation of *whiter*, *redder*

/wəɪt-ɻ/ → [wəɪɾ-ɻ]
/ɹɛd-ɻ/ → [ɹɛɾ-ɻ]

Plural allomorphy in English

cat [kæt] *cats* [kæt-s]
dog [dɑg] *dogs* [dɑg-z]
rose [ɹoʊz] *roses* [ɹoʊz-əz]

The realization of the plural suffix is different in different environments

[-s, -z, -əz] are allomorphs of the plural morpheme

Morphological rules for the English plural

plural ↔ /-s/ / voiceless consonants_
plural ↔ /-əz/ / {alveolar, postalveolar} fricatives_
plural ↔ /-z/ / elsewhere

X ↔ /Y/ / Z: 'X is realized as Y in environment Z'

Allomorphs

Phonologically distinct realizations of a morpheme

Morphemes can have multiple allomorphs which are in complementary distribution

e.g. Korean nominative case

pap-i 'rice-NOM' *sok-i* 'inside-NOM'
san-i 'mountain-NOM' *se-ka* 'bird-NOM'
pal-i 'foot-NOM' *pi-ka* 'rain-NOM'
waŋ-i 'king-NOM' *so-ka* cow-NOM

Korean nominative case rule

NOM ↔ /-i/ / C_

NOM ↔ /-ka/ / V_

Korean has two suppletive nominative suffixes

Suppletion

Suppletion is the alternation between two allomorphs of a morpheme which are not phonologically similar, the alternation between them is called

Suppletion in English

child	*child-ren*
ox	*ox-en*

-ren and *-en* are suppletive plurals in English

These allomorphs are lexically conditioned: they occur only with certain words or lexical items

e.g. Korean nominative case (cont.)

pap-i [pabi] 'rice-NOM'	*sok-i* [sogi] 'inside-NOM'
san-i [sani] 'mountain-NOM'	*se-ka* [sega] 'bird-NOM'
pal-i [paɾi] 'foot-NOM'	*pi-ka* [piga] 'rain-NOM'
waŋ-i [waŋi] 'king-NOM'	*so-ka* [soga] cow-NOM

Korean intervocalic voicing/rhotacization

/voiceless stops/ → [voiced stops] / V_V

/l/ → [ɾ] / V_V

Summary of analysis: two sequential processes

1. Allomorph selection, based on stem
2. Intervocalic voicing

Steps for morphological analysis

1. Organize the data, or carefully inspect the data and write down types of meanings you might need to account for
2. In words that share a piece of meaning for recurrent pieces of phonology, and hypothesize morpheme boundaries
3. Check for each morpheme in each words, revise
4. Identify the meaning which goes with the morpheme
5. Write down the morphemes and their meanings

Goals for morphological analysis

1. Identifying every morpheme and its meaning
2. Determining the orders in which morphemes occur
3. Identifying allomorphs of a single morpheme
4. Determining the distribution of allomorphs and writing their morphological rules

Morphologically conditioned stem allomorphy

walk	*walk-ed*
kiss	*kiss-ed*
buzz	*buzz-ed*
sway	*sway-ed*
kid	*kidd-ed*
lift	*lift-ed*

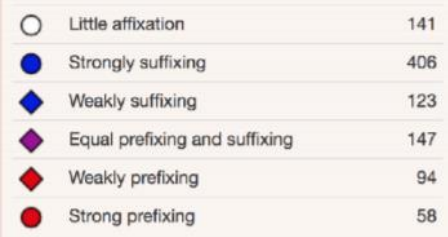
leave *left*
go *went*
The past tense suffix -ed has three allomorphs
The stems *leave* and *go* undergo suppletive allomorphy

Lecture 8.1: Morphology beyond prefixes and suffixes

Types of affixes

1. Prefixes PRE-Stem
2. Suffixes Stem-SUF
3. Circumfixes PRE-Stem-SUF
4. Infixes St-IN-em

English has both prefixes and suffixes, but more suffixes
Across languages, there is a tendency towards suffixation



(http://wals.info)

Circumfixes

e.g. past participles in German

	Infinitive	-past-1sg	Participle
'play'	spiel-en	spiel-t-e	ge-spiel-t
'work'	arbeit-en	arbeit-et-e	ge-arbeit-et
'say'	sag-en	sag-t-e	ge-sag-t
'hike'	wander-n	wander-t-e	ge-wander-t

ge- -t: an inflectional circumflex marking participial verbs

Infixes

e.g. infinitives in Tagalog

Verb Stem	Infinitive
lakad 'walk'	lumakad 'to walk'
bili 'buy'	bumuli 'to buy'
kuha 'take, get'	kumuha 'to take, to get'

-um-: an inflectional infix meaning 'to V'

The position of infixes is often phonologically specified (for Tagalog, after first C-)

English expletive infixation

1. ábso-fuckin-lútely
2. fán-fuckin-tástic
3. írre-fuckin-spónsible

-fuckin'-: an infix which expresses enthusiasm in casual speech

4. *áb-fuckin-solútely
5. *fantás-fuckin-tic

6. *írrespón-fuckin-siblé

-fuckin'-: must precede a stressed syllable, likes to go early in the word

Compounding

Compounding is the combination of open lexical categories to form a new word

(6) Examples of English compounds

Compounding of Free Morphemes	Compounding of Affixed Words	Compounding of Compounded Words
girlfriend	air-conditioner	lifeguard chair
blackbird	ironing board	aircraft carrier
textbook	watch-maker	life-insurance salesman

Compound stress

Compounds only receive one primary stress per word

e.g. income tax preparation fees

e.g. mint chocolate chip ice cream waffle cone

(8)

Compounds	*Phrases*
bláckbird	blác bírd
mákeup	máke úp

Reduplication

Reduplication is the repetition of a word, part of a word, or a phrase to mark a grammatical category

e.g. English phrasal reduplication

(11) Do you just like him as a friend, or do you like-like him?

(12) That shirt isn't what I had in mind; it's much too pale of a green. I want a shirt that is green-green.

(13) Yesterday we just went out for coffee, but this weekend we're going on a date-date.

e.g. Indonesian plural (total reduplication)

(14)

Singular		*Plural*	
rumah	'house'	rumahrumah	'houses'
ibu	'mother'	ibuibu	'mothers'
lalat	'fly'	lalatlalat	'flies'

e.g. Tagalog future tense (partial reduplication)

(15)

Verb Stem		*Future Tense*	
bili	'buy'	bibili	'will buy'
kain	'eat'	kakain	'will eat'
pasok	'enter'	papasok	'will enter'

Like infixation, partial reduplication has to be phonologically specified as a kind of rule or process

e.g. Tagalog occupation derivation (prefix + partial reduplication)

[mamimili]	'buyer'	< /maŋ+bi+bili/	[bili]	'buy'
[manunulat]	'writer'	< /maŋ+su+sulat/	[sulat]	'write'
[maŋʔiʔisda]	'fisherman'	< /maŋ+ʔi+ʔisda/	[ʔisda]	'fish'

Alternations

Alternations are morphological contrasts marked with a partial change to the stem

e.g. English plurals
mouse ~ mice
louse ~ lice
man ~ men
w[ʊ]m[ə]n ~ w[ɪ]m[ə]n
e.g. English tense inflection
ring ~ rang ~ rung
sing ~ sang ~ sung
drive ~ drove ~ driven
speak ~ spoke ~ spoken
e.g. Hebrew noun-verb derivation

Verbs		***Nouns***	
[limed]	'he taught'	[limud]	'lesson'
[sijem]	'he finished'	[sijum]	'end'
[tijel]	'he traveled'	[tijul]	'trip'
[bikeɣ]	'he visited'	[bikuɣ]	'visit (noun)'
[dibeɣ]	'he spoke'	[dibuɣ]	'speech'

Suppletion
Suppletions are morphological contrasts marked with a complete change to the stem
e.g. Suppletion in English

Present		***Past***	
[ɪz]	is	[wʌz]	was
[goʊ]	go	[wɛnt]	went

Adj		***Comparative***		***Superlative***	
[gʊd]	good	[bɛɾɨ]	better	[bɛst]	best
[bæd]	bad	[wɹs]	worse	[wɹst]	worst

e.g. Suppletion in Classical Arabic

Singular		***Plural***	
[marʔat]	'woman'	[nisaːʔ]	'women'

Singular		***Plural***	
[diraːsat]	'(a) study'	[diraːsaːt]	'studies'
[harakat]	'movement'	[harakaːt]	'movements'

Modeling alternations and suppletion as allomorphy
English plurals:
mouse *mice*(-PL)
man *men*(-PL)
English tense inflection:
ring-(PRES) *rang*-(PST)
sing-(PRES) *sang*-(PST)

<u>Lecture 8.2: Word Structure</u>
Three types of allomorphs
Allomorphs can be phonologically conditioned
kick-ed [kɪk-t] vs. rigg-ed [rɪg-d] vs. loot-ed [luɾ-əd]
Allomorphs can be morphologically conditioned
dorm-ir 'to sleep' vs. duerm-o 'I'm sleeping'

Allomorphs can be lexically conditioned

child-ren, ox-en

morpheme ↔ /allomorph/ / environment

'Underlying form is realized as surface form in environment'

Phonologically conditioned allomorphy

kick-ed /kɪk-t/	rigg-ed /rɪg-d/	loot-ed /lut-əd/
kiss-ed/kɪs-t/	buzz-ed /bʌz-d/	wad-ed /weɪd-əd/
	hir-ed /hɑɪɹ-d/	

Morphological rules:

past ↔ /-t/ / voiceless consonants_

past ↔ /-əd/ / alveolar stops_

past ↔ /-d/ / elsewhere

Morphologically conditioned allomorphy

duerm-o 'I sleep' (1st person singular)	dorm-imos 'we sleep' (1st person plural)
duerm-es 'you sleep' (2nd person singular)	dorm-eís 'you sleep' (2nd person plural)
duerm-e '(s)he sleeps' (3rd person singular)	duerm-en 'they sleep' (3rd person plural)

Morphological rules:

sleep ↔ /dorm/ / elsewhere

sleep ↔ /duerm/ / _-present, {sg, 3p}

Alternations as morphologically conditioned allomorphy

English plurals: *mouse* ~ *mice*-Ø

mouse ↔ /məɪs/ / _-plural

plural ↔ /-Ø / / *mouse*-_

English tense inflection: *ring* ~ *rang*-Ø

ring ↔ /ræŋ/ / _-past

past ↔ /-Ø/ / *ring*-_

mice and *rang* are morphologically conditioned allomorphs

/-Ø/ is a lexically conditioned allomorph (conditioned by roots/words)

Hierarchical word structure

A curious ambiguity:

unlockable: Incapable of being locked.

unlockable: Capable of being unlocked.

The explanation:

Morpheme order is not the whole story; complex words have internal structure

The structure of *reusable*

1. reusable, Adj.

   -able : V → Adj

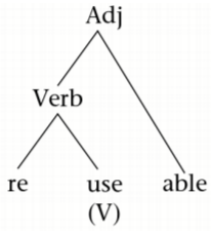
   re- : V → V

2. The correct derivation:

   use (V) → re-use (V) → [re-use]-able (Adj.)

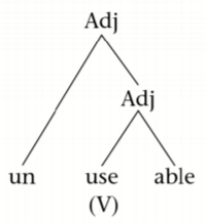
3. An impossible derivation:

   use (V) → [use-able] (Adj) → re-[use-able] (Adj.)



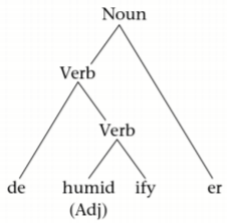
The structure of *unusable*

1. unusable, Adj.
   -able : V → Adj
   un- : V-to-V, Adj-to-Adj
2. The correct derivation:
   use (V) → use-able (Adj) → un-[use-able] (Adj.)
3. An impossible derivation:
   use (V) → *un-use (V) → [un-use]-able (Adj.)

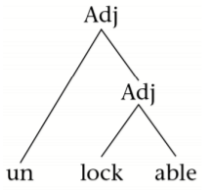


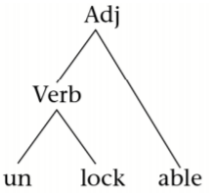
The structure of *dehumidifier*

1. de-humid-ifi-er, N.
   de- : V → V
   -ify : Adj → V
   re- : V → N
   Adj → V → V → N
2. The correct derivation:
   humid (Adj) → humid-ify (V) → de-[humid-ify] (V) → [de-[humid-ify]]-er (N)
3. Impossible derivations:
   humid (Adj) → de-humid (???); [humid-ify]-er (N) → de-[[humid-ify]-er]



The ambiguity of *unlockable*





Steps for solving word structure problems:

1. Identify the category of the entire word
2. Break the word into morphemes
3. Identify the root and its category
4. Determine which of the elements to its right or left the root combined with to form a new word
5. Repeat with the resulting complex word and repeat until you get to the top

<u>Lecture 8.3: Morphological Typology</u>

Morphological vs. phonological rules

Many instances of allomorphs can be modeled with either morphological or phonological rules. How do we know the difference?

Full productivity - a rule applies in all contexts (stem internally and across morpheme boundaries)

Partial productivity - a rule applies only within a particular morpheme

The Productivity/Regularity Criterion

1. Phonological rules are fully productive
2. Morphological rules are partially productive

If we have rule that a rule always applies, it should be modeled phonologically

If it only applies within a particular morpheme, it should be modeled morphologically

Morphological types

Analytic vs. synthetic languages

- Analytic language - 1 morpheme ↔ 1 word
- Synthetic language - multiple morphemes per word

The crucial difference between analytic and synthetic languages is the use they make of inflectional morphology

Analytic languages: Mandarin

- No subject agreement
- No tense inflection
- No inflectional plural

Do isolating languages "lack morphology"? No.

- All languages make some use of derivational morphology
- Almost all languages have some instances of compounding and reduplication
- Not all cases of derivational morphology look like affixation

e.g. V-V compounds in Mandarin

a. tamen **za-sui-**le yi-kuai boli.
   they pound-broken-LE a-CL glass
   'They smashed a piece of glass.'

b. wo **zhui-lei**-le ta le.
   I chase-tired-LE him SFP
   i. 'I chased him, which made him tired.'
   ii. 'I chased him, which made me tired.'

Synthetic languages

- e.g. case in Hungarian

(4) [ɔz ɛmber la:tjɔ ɔ kuca:t]
the man-(subject) sees the dog-(object)
'The man sees the dog'

(5) [ɔ kucɔ la:tjɔ ɔz ɛmbɛrt]
the dog sees the man-(object)
'The dog sees the man'

[ɔ ha:zunk zøld]
the house-our green
'Our house is green'

[ɔ se:kɛd ɔ ha:zunkbɔn vɔn]
the chair-your the house-our-in is
'Your chair is in our house'

Three types of synthetic languages

- Agglutinating language, e.g. Swahili
  - Easy to determine morpheme boundaries
  - Most morphemes express one feature or types of feature
- Fusional language, e.g. Spanish
  - Difficult to determine morpheme boundaries

Many morphemes express multiple features
Alternations/suppletions more common

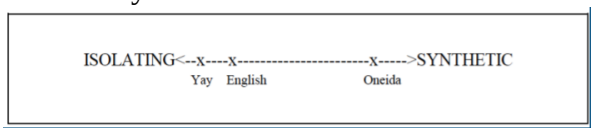
Polysynthetic language, e.g. Sora

Almost all arguments can be inflected on the verb
Can express meaning of an entire sentence with inflection and derivational morphology

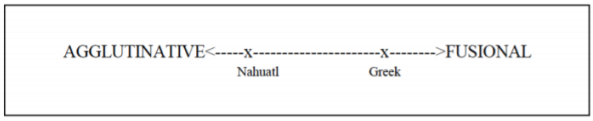
Morphological types as continuums

Morphological types are not discrete categories, but continuums
Some languages are 'more polysynthetic' or 'more analytic' than others
Within a single language, some morphemes can seem 'more fusional' or 'more agglutinating'

Index of synthesis



Index of fusion



Lecture 9.1: Child Language Acquisition (Berko 1958)

Two theories of child language acquisition
B.F. Skinner vs. Noam Chomsky
Two theories of child language acquisition

Behaviorism (B.F. Skinner): Humans are born *tabula rasa*, without any innate predispositions, including for language

Children lack any innate ability to learn language; humans are not special
Children learn (language) through repetition and reinforcement
A sentence that anyone utters at any point in time is a conditioned response to stimuli

Innatism (Noam Chomsky): Language is innate to *homo sapiens*, it is an evolved biologic trait; language learning is automatic

Human children are born with the innate (genetically coded) ability to learn language
The ability is in the form of a language acquisition device that allows them to hypothesize an internal mental grammar
A child's internal grammar allows them to generate novel utterances (and words) they have never heard before

What are children learning when they learn language?

Hypothesis 1: Children are memorizing sentences and words that they repeat
Hypothesis 2: Children are learning a grammatical system of rules that allows them to make sentences and words

The child's learning of English morphology (Berko 1958)

The ‘wug’ test
- Invented by Jean Berko (professor emeritus at Boston University)
- The first use of nonsense words to investigate children language acquisition
- Demonstrated that even young children have knowledge of morphological rules

Three groups of subjects
- Adult group (n = 12)
- Preschool group (n = 18, 4-5 years old)
- First grader group (n = 61, 5-7 years old)

Experimented tested four sets of rules
- English plural allomorphy /-z/, /-s/, /-əz/
- English past tense allomorphy /-d/, /-t/, /-əd/
- 3rd person sg. allomorphy /-z/, /-s/, /-əz/
- Possessive -’s allomorph

Also tested compound words

Were children using their linguistic abilities?
- “It was, moreover, evident that a great number of these children thought they were being taught new English words. It was not uncommon for a child to repeat the nonsense word immediately upon hearing it and before being asked any questions.”

Rating the responses
- Adults were unanimous except for /hif/ vs. /hifs/ ~ /hivz/

Sex-based differences
- Statistically, the 28 girls and 28 boys did equally well on all of the items
- No sex-based differences in morphology

Age-based differences
- First graders did significantly better than pre-schoolers on about half the items
- “During this period, children perfect their knowledge of these rules”

Results: the plural
- /wugz/ and /biks/ do not provide clear evidence for the morphological rule, */gs/, /kz/
- /n/, /l/, /r/, /w/, /j/ and vowels allow both /s/ or /z/: hens vs. hence

What mistakes did children make?
- With words like *tor* and *cra*, children never tried to apply the incorrect /-s/ form
- Instead, incorrect plural answers involved the repetition of the singular
- “The child here, as in so many other stages of language learning, answered complexity with silence”
- /hif/ vs. /hifs/ ~ /hivz/
  - Only two children opted for *heaves*
  - 10% of the children said /hijf-əz/, treating /f/ like the other (post-)alveolar affricates
- *glass* vs. *tass*
  - 91% of children could correctly supply the plural of *glass*
  - Only 36% could supply the form /tas-əz/
  - Only 26% could supply the correct plural of niz
  - “Again, the wrong answer consisted of doing nothing to the word as given.”

Compound words

"The object of this questioning was to see if children at this age are aware of the separate morphemes of a compound word"

Four kinds of explanation

1. Identity
2. Salient function or feature
3. Salient function connected to name
4. Etymological

"Private explanations" for compound words

"An airplane is called an airplane because it is a plain thing that goes in the air."

"Breakfast is called breakfast because you have to eat it fast when you rush to school."

"Thanksgiving is called that because people give things to one another."

"Friday is a day when you have fried fish."

"A handkerchief is a thing you hold in your hand, and you go 'kerchoo'."

These explanations show that speakers have access to the parts of compounds

What does this show?

Not just memorization

"If knowledge of English consisted of no more than the storing up of many memorized words, the child might be expected to refuse to answer our questions on the grounds that he had never before heard of a *wug, for instance, and could not possibly give us the plural form since no one had ever told him what it was. This was decidedly not the case."

Children often emphasized the endings in their answers

Children often pondered the right answer, or said "that's a hard one"

Three main findings

1. Children in the tested age range operate with clearly delimited morphological rules
2. Girls and boys did equally well
   "these results are at variance with the results of other language tests…girls have been shown to have a slight advantage over boys."
3. Morphological rules are still developing over this period; a simplified grammar was adopted by many children
   plural ↔ /-s/ / voiceless consonant_
   plural ↔ /-z/ / elsewhere

Did Berko settle the behaviorism vs. innatism debate?

Yes, at least partially

Children aren't just memorizing words and sentences

Children are learning systems of rules

On the other hand

Sufficiently powerful statistical models can often predict the right form without any explicit rule

Another question is whether these cognitive tools are necessarily exclusively for language as Chomsky claimed

Review: The Complexity of Language

Language is complex

Phonetics: The study of speech sounds

Descriptive vocabulary: fricative, stop, voiced, etc.
Universals: the IPA as possible speech sounds
Grammar of a particular language: phonetic inventories

Phonetics: The study of speech sounds

Phonology: The study of sound systems, i.e. sound grammar

Descriptive vocabulary: phoneme, allophone, syllable
Universals: syllable structure (CV), types of phonological rules
Grammar of a particular language: phonological rules
Natural classes as interface between phonology and phonetics

Phonetics: The study of speech sounds

Phonology: The study of sound systems, i.e. sound grammar

Morphology: The study of sound systems, i.e. word grammar

Descriptive vocabulary: derivation, inflection, morpheme, allomorph
Universals: all languages have complex words; complexity varies
Grammar of a particular language: morphemes, morphological rules
Phonologically-conditioned rules as interface between phonology and morphology

Phonetics: The study of speech sounds

Phonology: The study of sound systems, i.e. sound grammar

Morphology: The study of sound systems, i.e. word grammar

Syntax: The study of sentence grammar

Phonetics: The study of speech sounds

Phonology: The study of sound systems, i.e. sound grammar

Morphology: The study of sound systems, i.e. word grammar

Syntax: The study of sentence grammar

Semantics: The study of linguistic meaning

Lecture 10.1: Introduction to Syntax

Syntax and semantics

Basic notions

Grammatical - a sentence that is syntactically well-formed

1. Everybody surfs on their birthday.
2. My purple elbow annoyed the grumpy senator.

Ungrammatical - a sentence that is not syntactically well-formed

3. * Birthday the surfs on everybody.
4. * I really admire John's that painting.

Semantically anomalous - a sentence that is not semantically interpretable

5. # The shirt surfs lately.
6. # The purple senator annoyed my grumpy elbow.

Anomaly vs. ungrammaticality

Can you fix it just by replacing words with the same category but different meaning?

Semantic anomaly

Do you need to change the order of words, add or take away words, or change the inflectional morphology?

Syntactic ungrammaticality

This very clearly helps us understand the scope of syntax

Grammars and languages

Grammaticality is defined relative to a language

Grammaticality has nothing to do with prescriptive norms, how people "should" talk or write

On the contrary, every language and dialect has a complex set of internalized rules

Grammaticality reflects the judgments of native speakers of any particular language or dialect

Structure and meaning

Sentences can be grammatical but anomalous

1. # Colorless green ideas sleep furiously.

We can know that sentences are ungrammatical without knowing their meaning

2. * Green sleep colorless furiously ideas.

We can understand sentences that are ungrammatical

3. * Dog there biting I now ouch.

The independence of syntax and semantics

Two grammatically distinct sentences can have the same semantic content (= meaning)

1. The dog licked the child.
2. The child was licked by the dog.

These kinds of pairs tell us that meaning does not determine syntax

Syntax and structure

Sentences have structure

Syntactic patterns are structure dependent

1. * <u>Any</u>body does<u>n't</u> [like toddlers].

2. Toddlers don't [like anybody].
3. * [The toddlers I didn't meet] [like anybody].
4. [The toddlers I met] don't [like anybody].

Syntax is the theory of sentence structure

The syntax-semantics interface

Syntax builds complex structures

1. [The man in the park] [likes [the dogs]].
2. [The dogs in the park] [like [the man]].

Semantics interprets those structures

[Syntax] → [Semantics]

Goals for a theory of syntax

To predict what sentences in a language are grammatical and what sentences are ungrammatical, modeling a native speaker's knowledge of that language

Our formal theory of syntax should build complex syntactic structures

Our general theory should be applicable to any language

Establishing such a theory will hopefully help us understand better what is special about language and the human mind

Lecture 10.2: Syntactic Properties

Word order

Crucial to understanding the syntax of a language

One kind of syntactic property, among many others

When we talk about the word order of a language, we are interested as a starting point in basic word order

Word order in English

The elements in English sentences must occur in a particular order

1. a. Sally walked. Subject - Verb (SV)
   b. *Walked Sally. Verb - Subject (VS)
2. a. Sally ate an apple. Subject - Verb - Object (SVO)
   b. *Sally an apple ate. *Subject - Object - Verb (SOV)
   c. *Ate Sally an apple. *Verb - Subject - Object (VSO)
   d. *Ate an apple Sally. *Verb - Object - Subject (VOS)

Word order typology

Value	Representation
Subject-object-verb (SOV)	565
Subject-verb-object (SVO)	488
Verb-subject-object (VSO)	95
Verb-object-subject (VOS)	25
Object-verb-subject (OVS)	11
Object-subject-verb (OSV)	4
Lacking a dominant word order	189
Total:	1377

Co-occurrence requirements

If X occurs in a sentence, Y is required to occur before/after X

"If *the* occurs in a sentence, a *noun* is required to occur after *the*"

1. the dog
2. the cat
3. *the

"If *-able* occurs in a sentence, a *verb* is required to occur before *-able*"

1. abridge-able
2. walk-able
3. *-able

Arguments

If X necessitates the occurrence of Y, Y is an argument of X

e.g. noun phrases can be arguments of verbs

1. [The dog] chased [the cat].
2. *The dog chased.
3. *Chased the cat.

Omitting arguments from a sentence results in that sentence being ungrammatical

Subjects vs. complements

Two kinds of arguments of verbs

Subjects: an argument to the left of a verb

Complements: an argument to the right of a verb

Subjects are special

Only one subject can occur per sentence

Every sentence must have a subject

Verbs in many languages agree with subjects but not objects

Object agreement is possible, but only in languages that also have subject agreement

1. It/the dog like-s the cat.
2. They/the dogs like-Ø the cat.
3. It/the dog like-s the cats/them.

Arguments and flexible word order

e.g. Serbo-Croatian

Languages with flexible word order still show clear co-occurrence evidence for argumenthood

Different types of syntactic complements

(14) a. Sally told Polly she's leaving.
[*Polly* and *she's leaving* are both complements of *told*]

b. Sally put the book on the desk.
[*the book* and *on the desk* are both complements of *put*]

c. Sally persuaded Bob to go on vacation.
[*Bob* and *to go on vacation* are both complements of *persuaded*]

Not only verbs take arguments

e.g. Sally came to the part with Bob. (*with* takes a noun phrase complement)

e.g. Sally is fond of parties. (*fond* takes a prepositional phrase complement)

e.g. Bob invited Polly and Sally to the party. (conjunctions take conjuncts as arguments)

Verbs specify the number of arguments they take

Verbs taking one argument

1. [The dog] slept.
2. [The cat] meowed.

Verbs taking two arguments

3. [The dog] chased [the cat].
4. [The dog] likes [the cat].

Verbs taking three arguments

5. [The dog] took [the cat] [to the lady].
6. [The lady] gave [the dog] [a bone].

Adjuncts

There must be exactly the right number of arguments for all expressions in a sentence

At the same time, sentences contain a number of expressions which are optional called adjuncts

The presence or absence of adjuncts does not affect the grammaticality of the sentence

Adjectives in noun phrases are adjuncts

1. Sally likes dogs.
2. Sally likes small dogs.
3. Sally likes small fluffy dogs.
4. Sally likes small fluffy brown dogs.

Semantically, adjuncts like adjectives are modifiers, i.e. they modify the meaning of the element they attach to

It matters where adjuncts occur

Adjectives don't like to modify proper names and can't modify verbs

The same expression can be an adjective or adjunct

(25) a. Sally urged Bob to study French. [argument of *urged*]
b. Sally went to France to study French. [adjunct]

(26) a. Sally put the book on the desk. [argument of *put*]
b. Sally's cat was sleeping on the desk. [adjunct]

(27) a. Sally's cat seemed cute. [argument of *seemed*]
b. Sally has a cute cat. [adjunct]

(28) a. Sally behaved very carelessly. [argument of *behaved*]
b. Sally did her homework very carelessly. [adjunct]

Arguments	**Adjuncts**
Obligatory	Optional
Cannot have more than required	Can have as many as you like
Cannot be freely ordered w.r.t. one another	Can be freely ordered w.r.t. one another

Lecture 10.3: Constituency

Syntactic distribution
- Syntactic categories: words with the same syntactic (and morphological) distribution, i.e. parts of speech
- Position in the sentence, relative to other words
- Ability to take affixes (e.g. past tense /-ed/ for verbs)

Syntactic categories
- Syntactic categories include open lexical categories
  - Noun, verb, adjective, adverb
- Syntactic categories also include closed lexical categories, a.k.a. functional categories
  - Determiner, preposition, conjunction

Constituents: groups of words that syntactically act as one unit
- Also called phrases
- All arguments of verbs and adjuncts are constituents

Some basic types of constituents
- Noun phrase (NP): determiner + adjectives + noun
  - Noun phrases can be arguments of verbs
- Prepositional phrase (PP): preposition + noun phrase complement
  - Prepositional phrases are often adjuncts, so they can be repeated
- Sentence (S): a verb with all of its arguments and adjuncts
  - Sentences can also be complements of verbs
- Verb phrase (VP): verb + complements

Constituency tests: syntactic constructions/behaviors which provide evidence for the existence of constituents
1. Substitution test
2. Fragment answer test
3. Clefting/fronting test
4. Pseudocleft test

Substitution test: if a sequence of words can be replaced by a pro-form (like a pronoun), they form a constituent

Pro-NP: he/she/it/they (pronoun)
Pro-PP: here/there
Pro-VP: do so/did so
Pro-S: so

NP substitution

Ten penguins chased a boat in the middle of the desert.
They chased a boat in the middle of the desert.

PP substitution

They chased a boat in the middle of the desert.
They chased a boat there.

VP substitution

They chased a boat there.
They did so there.

S substitution

I said ten penguins chased a boat in the middle of the desert.
I said so.

Fragment answer test: make a question so that the answer is in the sequence of words in the question; if the words can be an answer, they are probably a constituent

NP fragment answer

Ten penguins chased a boat in the middle of the desert.
Q: Who chased a boat in the middle of the desert?
A: Ten penguins.

VP fragment answer

Ten penguins chased a boat in the middle of the desert.
Q: What did ten penguins do in the middle of the desert?
A: Chase a boat.

PP fragment answer

Ten penguins chased a boat in the middle of the desert.
Q: Where did ten penguins chase a boat?
A: In the middle of the desert.

Clefting/fronting test: try to put the words at the beginning of the sentence; if the sentence is grammatical, then they are probably a constituent

NP clefting/fronting test

I like this nice dog.
This nice dog, I like.

VP clefting/fronting test

Ten penguins chased a boat in the middle of the desert.
By golly, those ten penguins wanted to chase a boat somewhere...
and chase a boat, ten penguins did.

PP clefting/fronting test

Ten penguins chased a boat in the middle of the desert.
In the middle of the desert, ten penguins chased a boat.

S clefting/fronting test

I said ten penguins chased a boat in the middle of the desert.
Ten penguins chased a boat in the middle of the desert, I said.

Pseudocleft test: create a sentence with the following structure:
Wh-word - rest of sentence - was/is/were/are - constituent

NP pseudoclefting
Ten penguins chased a boat in the middle of the desert.
Who chased the boat were ten penguins in the middle of the desert.

VP pseudoclefting
Ten penguins chased a boat in the middle of the desert.
What ten penguins did was chase a boat.

PP pseudoclefting
Ten penguins chased a boat in the middle of the desert.
Where ten penguins chased a boat was in the middle of the desert.

Lecture 11.1: Syntactic categories

Lexical categories
- Lexical categories are the categories of individual words
- Open lexical categories
  - Noun, verb, adjective, adverb
- Closed lexical categories (or functional categories)
  - Determiner, preposition, conjunction
- Distinction between open and closed is whether they accept new members
- These seven categories will allow us to build a basic grammar

Identifying categories
- Syntactic distribution: the positions a category appears in
  - Notion of position can be defined by word order or structure
- Morphological distribution: the types of morphology, especially inflectional morphology, that a category can take

Morphological and syntactic distribution should converge on a single category

Determiners
- Syntactic distribution: must come before a noun and its adjuncts
- Only one determiner is allowed per noun phrase
- Articles (*the*, *a*)
  - The/a brown dog likes its toy.
- Demonstratives (*this*, *that*)
  - This/that brown dog likes its toy.
- Possessive pronouns (*his*, *her*, *their*, *my*, *your*, *our*)
  - His/her/their/my/your/our brown dog likes its toy.

Nouns
- Syntactic distribution: occur after determiners and adjectives
- Morphological distribution: can take plural -s
  - The brown dog likes its toy.
  - A brown dog likes its toy.

These/those brown dogs like their toys.

Adjectives

Syntactic distribution: occur between determiner and noun

Adjectives in noun phrases are adjuncts: they are optional, you can have as many as you want, and they freely order

Morphological distribution: can take comparative *-er*, superlative *-est*

The beautiful fluffy dog likes its toy.

That fluffy beautiful dog likes its toy.

Prepositions

Prepositions: *on*, *above*, *under*, *of*

Prepositions take a noun phrase complement

I put the mail on the table.

I put the mail on tables around the house.

*I put the mail on.

*I put the mail on happy.

Verbs

Syntactic distribution: occur between subject NP and object NP, sometimes after adverbs

Morphological distribution: can take past *-ed*, 3sg. present *-s*

The fluffy dog absolutely adores its chewy toy.

The tallest player really intimidated her opponents.

Subcategories of verbs

Verbs specify the number of arguments they take

Intransitive verb (= VP): no complement

[The dog] slept.

[The cat] meowed.

Transitive verb (= TV): one NP complement

[The dog] chased [the cat].

Ditransitive verb (= DTV): two NP complements

[The dog] brought [the lady] [the cat].

More subcategories of verbs

Verbs also specify the type of argument they take

Transitive verb (= TV): one NP complement

[The dog] chased [the cat].

[The dog] likes [the cat].

Sentential complement verb (= SV): one S complement

I think [the dog likes the cat].

I realized [the lady gave the dog a bone].

Adverbs

Modify the time, manner, or place of a verb phrase

Syntactic distribution: attach to VP; occur either before or after the verb and its complements

Morphological distribution: many adverbs end with *-ly*

The dog aggressively [chased the cat].

The dog [chased the cat] aggressively.
My sister already [gave a present to my mother].
My sister [gave a present to my mother] already.

Summary: Lexical categories

Nouns (N): words like *dog, thought*
Verbs (V): words like *run, sleep*
Adjectives (Adj): words like *blue, large*
Adverbs (Adv): words like *easily, well*
Determiners (Det): words like *a, the*
Prepositions (P): words like *under, on*

Phrasal categories

Constituents are phrasal categories
Like lexical categories, phrasal categories have a particular syntactic distribution
Noun phrase (NP), verb phrase (VP), prepositional phrase (PP), sentence (S)

Noun phrase (NP)

Syntactic distribution: before verbs, as subjects, and immediately after transitive verbs and prepositions

[The fluffy dog]$_{NP}$ barked.
I preferred [the fluffy dog]$_{NP}$.
I walked with [the fluffy dog]$_{NP}$.

Prepositional phrase (PP)

Syntactic distribution: adjuncts occurring after nouns and adjuncts occurring after verbs

The bench [in the park]$_{PP}$.
The dog barked [in the park]$_{PP}$.

Verb phrase (PP)

Example: intransitive verb(IV) or transitive or ditransitive verb with its complement
Syntactic distribution: between subjects and verbal adjuncts such as adverbs

The fluffy dog [barked]$_{VP}$ loudly.
The fluffy dog [bit my leg]$_{VP}$ hard.
The fluffy dog [gave the lady a bone]$_{VP}$ yesterday.

Sentence (S)

Can occur as independent utterances or as the complement of sentential verbs

[The dog bit my leg]$_{S}$.
I claimed [the dog bit my leg]$_{S}$.

The structure of NP

Simple noun phrases

Some words have the distribution of an NP: pronouns, names

She slept.
Mary slept.

We can analyze these words as being of the category NP
We can thus introduce a phrase structure rule for these words

NP → she
NP → Mary

Phrase structure rules

Phrase structure rules are instructions for drawing trees

Trees are models of syntactic structure

Three components of noun phrases

Arguments: singular count nouns require a determiner

Adjuncts: adjectives can precede the noun

Noun: the head of the noun phrase

The presence of a noun is obligatory in a noun phrase

Lecture 11.2: Constructing a grammar

Review: Syntax

1. Sentences have internal structure
2. We an identify the structure using constituency tests
3. Some constituents are arguments, others are adjuncts
4. We can identify the category of words and constituents by examining their syntactic (and morphological) distribution
5. We also need to account for the order of words

Review: Lexical and phrasal categories

Open lexical categories:

- Noun (N), verb (V), adjective (Adj), adverb (Adv)

Closed lexical categories:

- Determiner (Det), prepositions (P), conjunctions (Conj)

Phrasal categories:

- Sentence (S), noun phrase (NP), verb phrase (VP), prepositional phrase (PP)

Goal of constructing a grammar

Generative grammars: an explicit procedure for constructing all grammatical sentences in a language – and no ungrammatical sentences

Speakers can generate new grammatical sentences (but not ungrammatical sentences), so this models mental grammar

This is a theory of competence, i.e. human knowledge of language, not performance, i.e. what we are doing when we speak

Limitations: it will be impossible to do this for all of English in this class, so we model a fragment of English grammar; what we won't model: auxiliary verbs, quantifiers, negation

Properties of our grammar

Argument rule for NP

Adjunct rule for N

Adjective optionality

Recursion (adjectives, sentence complements)

Adjective free ordering

VP structure

Adjunct properties

Adjuncts are optional

Adjuncts can occur as many times as you want

Adjuncts are freely ordered

Subcategories of verbs

Intransitive verb (= VP): no complement

[The dog] slept.

Transitive verb (= TV): one NP complement

[The dog] chased [the cat].

Ditransitive verb (= DTV): two NP complements

[The dog] brought [the lady] [the cat].

Sentential complement verb (= SV): one S complement

I think [the dog likes the cat].

Grammar so far

1. S → NP VP
2. NP → Det N
3. N → Adj N
4. VP → V_0
5. VP → V_1 NP
6. VP → V_2 NP NP
7. VP → V_S S

These seven rules have the following properties:

– They generate accurate representations of syntactic structure
– They capture the distinction between adjuncts and arguments
– They construct only grammatical sentences, and predict the distribution of all the categories
– They allow us to construct sentences of infinite length

Sentences of infinite length

These illustrate the recursive property of natural language syntax

This property requires us to posit a grammar which can make sentences of infinite length

Adverbs as VP adjuncts

Adverbs are optional (optionality)

The dog (aggressively) [chased the cat] (aggressively).

Adverbs occur multiple times (recursive)

The dog [chased the cat] aggressively already

Adverbs are freely ordered (free ordering)

The dog [chased the cat] already aggressively

<u>Lecture 11.3: Constructing a grammar (part 2)</u>

Prepositions

Examples: *on*, *above*, *under*, *of*

Prepositions take a noun phrase complement

The distribution of *of* is different, we set it aside

Prepositional phrase (PP)

Adjunct occurring after NPs and VPs
PPs include a preposition and an NP
Prepositions have NP complements
PPs are adjuncts that can attach after nouns/NPs and VPs
Recursive

Two types of ambiguity

Lexical ambiguity (homophony)

When different words happen to be pronounced the same
Ambiguous words can differ in their lexical category
e.g. reader

a. We should find some papers and collect them into a reader.
b. Sally is an avid reader of science fiction.

e.g. that

c. Sally likes that.
d. Sally likes that dog.

Syntactic ambiguity

When sentences have multiple meanings, but not due to any ambiguous words
e.g. The cop watched the man with the binoculars.
e.g. Sandy said Tom would be here yesterday.

Making ambiguity disappear

Analyzing ambiguity with constituency tests

Comments on ambiguity

Syntactic ambiguity shows that word order ≠ syntactic structure

Linear order can be predicted from a single syntactic structure
The syntactic structure can't be predicted from a string of words

Syntactic ambiguity shows that sentence meaning depends on syntactic structure

Sentence meaning can be predicted from a syntactic structure
Syntactic structure can't be predicted from sentence meaning

Is language perfect?

No; syntactic ambiguity is a terrible design feature of language
Infinite sentence length is a useless feature of syntax
Both of these properties are simple consequences of the fact that syntax is two-dimensional (ambiguity) and recursive (infinite sentences)

Structure and recursion are basic properties of language
They make syntax possible, which makes language possible

Language isn't "designed" as a perfect system

Grammar so far

1. S → NP VP
2. NP → Det N
3. NP → Pron
4. N → Adj N
5. N → N PP
6. PP → P NP

7. VP → V_0
8. VP → V_1 NP
9. VP → V_2 NP NP
10. VP → V_S S
11. VP → Adv VP
12. VP → VP Adv
13. VP → VP PP

Lecture 12.1: Drawing trees

Coordination

Coordinating conjunctions (and)

e.g. The tall lady and I admired the wombat.

e.g. The little boy saw the wombat and screamed.

Coordinating disjunctions (or)

e.g. The wombat slept in my bed or under the table.

Properties of coordination

Conjunctions are lexical items that take two arguments

Both arguments must be of the same category

The phrase resulting from conjunction behaves like the arguments of the conjunction

Recursive and results in more syntactic ambiguity

Final grammar for this course

1. S → NP VP
2. NP → Det N
3. NP → Pron
4. N → Adj N
5. N → N PP
6. PP → P NP
7. VP → V_0
8. VP → V_1 NP
9. VP → V_2 NP NP
10. VP → V_S S
11. VP → Adv VP
12. VP → VP Adv
13. VP → VP PP
14. NP → NP Conj NP
15. VP → VP Conj VP
16. PP → PP Conj PP
17. S → S Conj S

How to draw trees

Rules of thumb for drawing trees

1. Don't cross lines
2. Begin by identifying lexical categories

3. Make sure that you are following your phrase structure rules
4. Double check your constituents

A general strategy: bottom-up

1. Identify all of your lexical categories, paying particular attention to the verb
2. Underline the subject NP (the rest is VP)
3. Identify the phrase structure rules that you're going to need for the sentence
4. Start with the verb, applying rules to VP, then combine VP adjuncts, then look for the subject NP, then go to S
5. If you have an S complement verb, start with the embedded sentence

<u>Lecture 12.2: Japanese syntax</u>

Japanese syntax: an overview

Word order is SOV (sometimes)
Postpositions
Embedded clauses precede verbs (like objects)
Adverbs precede verbs

The verb-final constraint

1. John-ga Mary-o mi-ta — SOV
   John-NOM Mary-ACC see-PAST
   'John saw Mary'
2. Mary-o John-ga mi-ta — OSV
3. John-ga mi-ta Mary-o — *SVO
4. Mary-o mi-ta John-ga — *OVS

Some general observations

Languages with SOV order tend to order OSV as an alternate word order
Languages with SOV order tend to have case markers
Languages with SOV order tend to be postpositional

Japanese syntax

Postpositions
    These are the translational equivalents of English prepositions
    They don't occur with case markers
Embedded clauses
    Embedded clauses precede verbs
Adverbs
    Adverbs precede verbs

Constructing a grammar for Japanese

Challenge: Try to keep the rules as similar to English as possible

Japanese grammar

1. $S \rightarrow NP\ VP$
2. $NP \rightarrow N$
3. $PP \rightarrow NP\ P$
4. $VP \rightarrow NP\ V_1$
5. $VP \rightarrow S\ V_S$

6. VP → PP VP
7. VP → Adv VP

Is syntactic structure universal?

- There are no major obstacles to extending our syntactic analysis of English to Japanese
- Extensive research has confirmed that the structural analysis just described for Japanese is essentially correct
- Despite having fundamental differences in their word order, Japanese and English have essentially the same syntactic structure

<u>Lecture 13.1: Headedness and Movement</u>

Headedness

- The category that labels a phrase is its head
  - N is the head of NP
  - P is the head of PP
  - V is the head of VP
  - (S has no head, as afar as we have been able to tell)
- Heads are obligatory and determine the distribution of the phrase
- Heads bear inflection relevant for that kind of phrase

Japanese rules

1. PP → NP P
2. VP → NP V_1
3. VP → S V_S

English rules

1. PP → P NP
2. VP → V_1 NP
3. VP → V_S S

Japanese vs. English

- Heads are on the right in Japanese
- Heads are on the left in English

Word order typology

- Postpositional languages are almost all SOV
- Languages where embedded S precedes verbs are all SOV
- In other words, languages tend to be left-headed or right-headed as a property of the entire language

The seeds of a theory

- Suppose that English & Japanese have the same phrase structure rules in an abstract sense
- The only difference is that in Japanese, heads go on the right and in English, heads go on the left (right-embedding in English vs. left-embedding in Japanese)
- This makes learning syntax for children easier

Summary

- Major word order differences between languages can be captured by proposing different phrase structure rules

Phrase structure rules in the same language tend to follow the same pattern, with heads at the beginning (e.g. English) or at the end (e.g. Japanese)

Other phrase structure rules might create other differences

Movement

Our grammar fragment so far includes:

A theory of categories (N, V, Adj, ...)

A theory of structure (phrase structure rules)

This theory is incomplete

Topicalization

e.g. (My cat doesn't usually eat mice but) <u>this mouse</u>, it [enjoyed __]$_{VP}$.

e.g. (I don't like most books but) <u>this novel</u>, I [finished __ quickly]$_{VP}$.

The NP/PP at the beginning of the sentence belongs in the VP

Wh-questions

Many other types of sentences behave similarly

e.g. <u>What</u> did the cat [eat __]$_{VP}$?

e.g. <u>Which book</u> did you [finish __ quickly]$_{VP}$?

e.g. <u>How</u> does Mary [eat soup __]$_{VP}$?

The wh-word at the beginning of the sentence belongs in the VP

Transformational grammar (Chomsky)

Syntactic structure is generated in two stages

1. Deep structure (phrase structure rules)
2. Surface structure

Topicalization rule: After completing the phrase structure rules, move any NP or PP to the front of a sentence as a topic

1. The cat ate <u>the rat</u>. Deep structure
2. <u>The rat</u>, the cat ate __. Surface structure

Wh-movement rule: After completing the phrase structure rules, move any wh-phrase to form a question

1. You read <u>which book</u>. Deep structure
2. <u>Which book</u> did you read __? Surface structure

Alternatives

The idea of "movements" and "derivations" is among the most controversial of Chomsky's proposals (how do we do this in our brains?)

Yet, there are few good alternatives:

1. Completely different phrase structure rules for questions/topicalization?
2. Silent NPs that must occur whenever there is a word initial topic/wh-phrase?

Grammar and processing

Relative clauses

Relative clauses involve a movement transformation that takes the subject out of a clause and makes it a separate NP

When relative clauses modify the objects, this creates recursion

NP [$_S$ ... V NP [$_S$ ... V NP [$_S$ ...

English relative clauses occur after the noun they modify

e.g. the cat [that __ ate the rat]

e.g. (English right-embedding)

John owned a cat that killed a rat that ate cheese that was rotten.

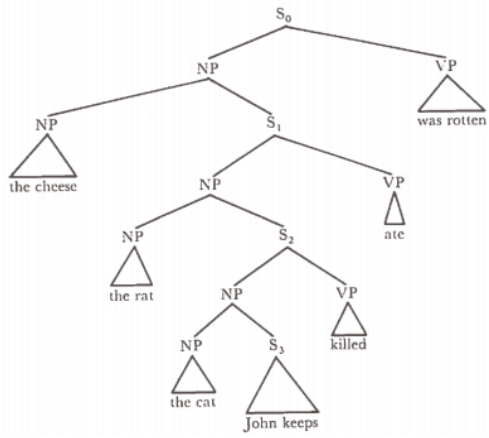
Japanese relative clauses occur before the noun they modify

e.g. (Japanese left-embedding)

[[[John ga katte-iru]$_S$ neko ga korosita]$_S$
nom. keep cat nom. killed
nezumi ga tabeta]$_S$ tiizu wa kusatte-ita.
rat nom. ate cheese theme rotten-was
(Literal word order) [[[John keeps] cat killed] rat ate] cheese was rotten.'

e.g. (English center-embedding)

The cheese that the rat that the cat that John keeps killed ate was rotten.



English and Japanese both have grammars with freely recursive relative clauses

Yet some instances of recursion are nearly impossible to understand

Therefore, our ability to use/understand a particular structure is distinct from the well-formedness of that structure according to the grammar

Competence vs. performance

Competence: native speakers' knowledge of grammatical rules

Performance: the production and comprehension of sentences by the brain

Chomsky claimed that a theory of competence must be distinct from a theory of performance

For Chomsky, phrase structure rules and transformations are not necessarily a theory of how your brain creates sentences

They represent an abstract theory of competence, which must be supplemented with a theory of performance

This is among Chomsky's most controversial claims

Lecture 13.2: Introduction to Semantics

What syntax told us about semantics

Syntactic ungrammaticality ≠ semantic anomaly

Sentence meaning is dependent on syntactic structure

Syntactic ambiguity provides evidence for this

A grammar for well-formed syntactic representations can be stated without reference to meaning

The Saussurean sign (Ferdinand de Saussure)

Signifier - /boʊt/

Signified - (the actual boat)

What makes a sign?

Linguistic signs are arbitrary

Linguistic signs are conventionalized

Linguistic signs have the capacity to be combined

Sense and reference

Two components to linguistic meaning

The reference of an NP is the person, object, or entity that it indicates

The sense of an NP is the "mode of presentation" the NP provides for that reference

Sense and reference capture distinct, but equally important aspects of linguistic meaning

The following NPs all have the same reference:

1. I visited China.
2. I visited that country. (pointing to China on a map)
3. I visited the most populous country in the world.
4. I visited the host country of the 2008 Summer Olympic Games.

If we did without reference...

...we couldn't explain how we can use language to talk about the world around us

...we couldn't explain the fact that the four sentences earlier are true in exactly the same situations

...we couldn't explain the difference between somebody who can distinguish real vs. fake diamonds and somebody who knows the meaning of the word 'diamond'

...we couldn't explain the difference in meaning between 'that horse' (possible) and 'that unicorn' (impossible)

If we did without sense...

...we would be unable to distinguish the following sentences:

1. The morning star is the evening star.
2. The evening star is the morning star.

...we would be unable to distinguish the four sentences from above

...we couldn't explain why the following four sentences:

1. I want to visit China.
2. I want to visit that country. (pointing to China on a map)
3. I want to visit the most populous country in the world.
4. I want to visit the host country of the 2008 Summer Olympic Games.

Types of linguistic meanings

Four kinds of meanings

Proposition: meaning of S

Reference: meaning of NP

Property: meaning of V_0, non-relational N, regular A
Relation: meaning of V_1, relational N, comparative A

Proposition: the claim that a sentence makes about the world
Sentences express propositions
Propositions have the capacity to be either true or false
The following sentences both express the proposition that Sally visited Molly
1. Sally visited Molly.
2. Molly was visited by Sally.

Reference
Referential expressions pick out individuals in the world
Only NPs refer; the canonical referring expressions are names and pronouns
Names are context independent
Pronouns are context dependent
I - whoever is talking at any particular moment
you - whoever I'm talking to at any particular moment
they - whoever I'm talking about at any particular moment

Property
Properties describe traits or attributes that hold of individuals
All open lexical categories (nouns, verbs, adjectives) can describe properties
Noun properties: *universe*, *dog*, *unicorn*, ...
Verb properties: *smile*, *sleep*, *run*, ...
Adjective properties: *tall*, *disgusting*, *blue*, ...

Relation
Relations describe kinds of relationships that hold between two individuals
All open lexical categories (nouns, verbs, adjectives) can describe relations
Noun relations: *mother*, *enemy*, *mayor*, ...
Verb relations: *hug*, *drink*, *admire*, ...
Adjective relations: *taller*, *grosser*, *better*, ...

<u>Lecture 13.3: Compositional Semantics</u>

Property sense vs. reference
The sense of a property is some kind of abstract concept which is difficult to characterize
Three ideas:
1. Property senses correspond to definitions
2. Property senses correspond to mental representations
3. Property senses correspond to uses

Yet properties 'refer' to sets of individuals

Issues with these conceptions of property senses
Property senses as definitions
e.g. *glass* can mean 'a receptacle for drinking made of glass'
Problems: dependent on language, circular, often imprecise or insufficient
Property senses as mental images
e.g. *glass* means what is in your brain when you hear 'bird'

Problems: the prototype problem, many words lack mental images (e.g. *universe*, *impossible*, *lie*)

Property senses as uses

e.g. *birds* are things that you use the word *bird* to talk about

Problems: meanings cannot exist independent of their uses, using *bird* to talk about a dog doesn't make it a bird

Properties as sets of individuals

e.g. *woman* is

Problems: we don't need to know every possible individual to use the word, what about nouns like *unicorn*?

The set-based meaning of properties is the most widely assumed

It easily accounts of sentences like *Mary is a woman.*

It easily can be expanded to include relational meanings, which are just sets of pairs/tuples, e.g. *mother* = {(Michelle, Sasha), (Michelle, Malia), ...}

It is compatible with externalism, the philosophical stance that the notion of truth is determined by the facts which hold in the world

Propositions

Sentence meanings

Sentences express propositions or claims about the world

Propositions have truth values, either true or false

To know the truth value of a sentence, you must know its truth conditions

If the truth conditions hold of the world, the sentence is true

If the truth conditions do not hold of the world, the sentence is false

Sentence sense and reference

The sense of a sentence is its truth conditions, i.e. the conditions in which it is true

The reference of a sentence is its truth value, true or false

Entailment: sentence X entails sentence Y if whenever X is true, Y must be true as well

Our language ability includes knowledge of entailment relations

Entailment is at the core of the human meaning and reasoning

e.g. 1. All dogs bark.

2. Sally's dog barks.

Sentence 1 entails sentence 2 because if it is true that all dogs bark, it must be true that Sally's dog barks

Sentence 2 does not entail sentence 1 because it could be true that Sally's dog barks without it being true that all dogs bark

The truth conditions of sentence 1 include the truth conditions of sentence 2

Entailment is a very strong relation: if X is true, Y must be true

Entailment ≠ truth values

We cannot know if X entails Y just be knowing truth values, only truth conditions

Entailment and reference

The reference of proper names is assumed to be part of one's knowledge of language, so we have entailment in the following pairs:

1. Ian has visited Spain.
2. Ian has visited Europe.

Semantic composition

What determines truth conditions?

1. The dog likes Sally.
2. Sally likes the dog.

These sentences have exactly the same three words, but their truth conditions are independent

Two factors in determining sentence meaning

The words in the sentence (lexical semantics)

The way the words are put together (syntax)

Principle of Compositionality: the meaning of a sentence (or any other multi-word expression) is a function of the meanings of the words it contains and the way in which these words are syntactically combined

Infinite syntax and semantics

Our syntactic knowledge allows us to build an infinite number of sentences

We have a corresponding semantic system which allows us to understand an infinite number of sentences, including those never heard before

Idioms

Some kinds of words and sentences have non-compositional meanings

e.g. 1. Polly <u>kicked the bucket</u>.
2. Polly <u>died</u>.

e.g. 1. Polly <u>gave</u> Tim <u>a hard time</u>.
2. Polly <u>teased</u> Tim.

Interpreting S → NP VP

How do we interpret the sentence, *Sandy runs*?

Sandy is referential, it picks out a specific individual

runs is a property, it describes traits which hold of a particular set of individuals

For *Sandy runs* to be true, it must be the case that Sandy is a member of the set of individuals that runs

<u>Lecture 14.1: Theta Roles</u>

Semantically anomalous sentences

1. # The wall kicked Francesca.
2. Francesca kicked the wall.
3. Francesca sleeps well.

4. # Colorless green ideas sleep well.
5. # The party surprised some driftwood.
6. The party surprised Francesca.

Entailments of verbs

1. Francesca kicked the wall.
2. Francesca's foot made forceful contact with the wall.
3. ~~Francesca doesn't like the wall.~~
4. Francesca moved her foot on purpose.
5. ~~The wall was injured.~~

Verb meaning

Entailments of: X kicked Y

X's foot made forceful contact with Y

X moved X's foot on purpose

Verbal entailments seem to be dependent on and make claims about their arguments

Certain verbal entailments about arguments occur across many different verbs and allow us to classify verbs

These entailments are called theta roles

Theta roles

We say that a verb assigns theta roles to their arguments

The number of theta roles a verb assigns corresponds to its number of arguments

e.g. *kick* assigns two theta roles because it has two arguments, an agent and a patient

Every NP must receive a theta role

Verbs must assign their theta roles to only one argument

NPs which are not arguments of verbs get assigned a theta role by their preposition

Seven theta roles

1. Agent
2. Experiencer
3. Patient
4. Instrument
5. Source
6. Goal
7. Location

Agents

In control of the action

Apply force

Cause the action

Often show an animacy restriction

1. <u>Francesca</u> jumped.
2. # <u>The table</u> jumped.
3. <u>Francesca</u> ran.
4. # <u>The table</u> ran.

Patients

Not in control of the action

Don’t apply force
Don’t cause the event
Are affected by the event, either in change of state or location

1. The table broke.
2. The package disappeared.
3. The ice melted.

Agents and patients

When agents and patients are in the same sentence in transitive verbs, the agent is always the subject

1. Francesca kicked *the wall*.
2. Ahmed kissed *Bernard*.
3. The elephant crushed *the tent*.
4. My hand melted *the ice*.

In passive sentences, the patient is the subject and the agent is in a by-PP, an adjunct

5. *The tent* was crushed by the elephant.

Experiencers

Must have a mental state
The verb describes the mental state of this argument
The mental state can be positive or negative
Always shows an animacy restriction

1. Fatima worried.
2. # The table worried.
3. Fatima admired Francesca.
4. # The table admired Francesca.

Like agents, experiencers can be subjects of intransitive or transitive verbs
Patients can only be subjects if the verb is intransitive
Experiencer objects

5. The storm worried Fatima.
6. Francesca annoyed Fatima.

Instruments

Describes the means by which an agent does the event
Usually adjuncts, introduced by *with* or *by*

1. Fatima jumped with the trampoline.
2. Francesca ate the sushi with a fork.

Instruments as subjects

3. The key unlocked the door.
4. The knife cut the bread.

Instruments as objects

5. The lady used the knife.

Sources

Describe the origin of motion or force
Usually introduced by prepositions (*from*, *out of*)

1. Fatima ran from the woods.

Goals

Describe the destination of motion or transfer

Often introduced by prepositions (*to*, *into*, *at*, *onto*, etc.)

1. Francesca jumped into the pool.

Goals in ditransitives

2. Fatima threw the dog a bone.
3. Francesca gave me a sticker.

Verbs with two NP complements (V_2) always assign a goal and a patient

The goal always comes first

Sources and goals

Goals and sources are often adjuncts of intransitive or transitive verbs

Because they are adjuncts, goals and sources occur together

1. A bug crawled out of the hole towards *the chef*.

Locations

Describe where the event takes place

Motion events must be contained within the location

Always introduced by prepositions, usually adjuncts

1. Fatima ran in the woods.
2. Francesca jumped on the moon.
3. Fred ate his dinner beneath the ocean.

Seven theta roles

Arguments of intransitive and transitive verbs

Agent, patient, experiencer, instrument

Adjuncts

Goal, source, location

Linking subjects/objects and theta roles

Why do certain theta roles get linked with certain positions?

Subjects of intransitive verbs can be:

Agent, patient, experiencer

Subjects of transitive verbs can be:

Agent, experiencer, instrument

Subjects of ditransitive verbs can be:

Agent

Complements are:

Patient, experiencer, goal

One solution: the theta hierarchy

Agent < Experiencer < Instrument < Patient < Goal < Source/Location

If a verb has more than one argument:

1. Assign the highest-ranked argument to *subject*
2. Assign the second-highest argument to *direct object*
3. Assign the third-highest argument to *indirect object*

Summary

One component of compositional meaning is that verbs introduce entailments about their arguments
These entailments fall into systematic classes called theta roles
Theta roles and the theta hierarchy help us understand why certain arguments are associated with certain meanings

Lecture 15.1: Cross-linguistic semantics

Principle of Compositionality: the meaning of a sentence (or any other multi-word expression) is a function of the meanings of the words it contains and the way in which these words are syntactically combined

Interpreting S → NP VP

How do we interpret the sentence, *Sandy runs*?
Sandy is referential, it picks out a specific individual
runs is a property, it describes traits which hold of a particular set of individuals
For *Sandy runs* to be true, it must be the case that Sandy is a member of the set of individuals that runs

Categories and meaning

Open lexical categories (N, V, Adj) can be properties or relations
Phrases can be referential (NP) or propositional (S)
We have not yet examine the meaning of closed lexical categories (Det, P, Conj)

The meaning of *the*

Sentences with and without *the*

1. I like marmots.
2. I like the marmots.
3. Marmots are cute.
4. The marmots are cute.

The meaning of *the*

How do we get from properties to reference?
marmot is a property, it describes traits which hold of a particular set of individuals
the marmot can refer, it refers to a particular individual like a pronoun or name
the is a function from properties to individuals

property → reference

Conditions on use

e.g. # The marmot is big. (when there are multiple marmots)
In order to use *the* to refer to a marmot, there must be a unique marmot in the context
The uniqueness condition on the use of *the* is called a presupposition
Many determiners do not carry a uniqueness presupposition

1. # The marmot is big.
2. A marmot is big.
3. Some marmot is big.
4. One marmot is big.

Two components in the meaning of definite articles like *the*

The referential function

the converts properties into referential expressions/individuals

The definite function

the requires uniqueness in the context

In contrast, indefinite articles like *a(n)* can be referential but never require uniqueness

Is meaning universal?

Universal: the same in every language

Given our ability to translate, it is appealing to think that meaning is actually the same in every language in some sense

Four sub-questions about semantic universality

1. Does every language have open lexical categories with the same meanings?
2. Does every language have closed lexical categories with the same meanings?
3. Does every language use the same basic mechanisms to construct sentence meanings?
4. Does every language have the exact same truth conditions?

Variation in open lexical items

It is not hard to find words in one language that are missing in another

e.g. Thai: kreŋcai (Adj) - 'reluctant to impose on someone'

e.g. Moro: umurtáɲ (N) - 'my co-spouse'

Languages without *the*

50% of the world's languages have no translation for *the*, e.g. Mandarin

Variation in the meaning of *the*

Familiarity

A problem for uniqueness?

1. A man walked into a room full of people.
2. The man was tall and menacing.

The uniqueness view (Gotlob Frege):

the (and definiteness) is about uniqueness

The familiarity view (Irene Heim):

the (and definiteness) is about prior use or knowledge of the referent

Definite articles in other languages

e.g. Lakota has two definite articles, a unique definite article and a familiar definite article

e.g. Familiar definites in Mandarin require a demonstrative determiner and bare nouns are only allowed in unique definite contexts, so Mandarin actually distinguishes different types of definiteness

Languages with a single definite article

In languages with a single definite article, like English, there may be ambiguity between the two meanings

Unique use of *the*

1. the Campanile
2. # a Campanile

Familiar use of *the*

1. A man walked into a room full of people.

2. The man was tall and menacing.

The definite article in English may be ambiguous between the two meanings

Alternatively, just one of these meanings might be sufficient for both contexts somehow

What does this mean?

1. Does every language have open lexical categories with the same meanings? No.
2. Does every language have closed lexical categories with the same meanings? No, but...

    Closed lexical categories are used to express meanings in particular grammatical contexts (like uniqueness and familiarity)

    The same limited set of grammatical meanings seem to be important to all languages (even if they don't have words for all of them)

Summary

The meanings of closed lexical categories are complex and require an understanding of compositionality

The meanings of open lexical categories vary widely, but can be paraphrased

The types of meanings that closed categories can make reference to seem to be roughly the same in all languages

<u>Lecture 15.2: Quantifiers</u>

Bach, et al. (1995): "Every natural language provides some means for making general statements."

What is quantification?

"Speaking notionally, quantification always has to do with some domain of entities about which some claim is made as to its relation to other domains."

In other words, quantification is a relation between properties

It is an additional type of meaning from individuals, properties, and relations

Quantificational sentences

1. Birds fly.
2. All birds fly.
3. Birds always fly.
4. Generally, birds fly.
5. Birds can fly.
6. The typical bird flies.
7. Usually, if something is a bird, it can fly.

Two types of quantification

Quantificational determiners (= D-quantifiers)

All birds fly.

Most birds fly.

Some birds fly.

No birds fly.

Quantificational adverbs (= A-quantifiers)

Birds always fly.

Birds mostly fly.

Birds sometimes fly.

Birds never fly.