Introduction to English Language & Linguistics

0. Introduction to language and linguistics

- 0.1. grammar = linguistics from school
- 0.2. linguistics = thinking about language
- 0.3. features of human language

1. Phonetics & phonology

2. Morphology & word formation

3. Syntax and grammar

4. Semantics, pragmatics and lexicology

5. Macrolinguistics

- 5.1. Textlinguistcs
- 5.2. Sociolinguistics
- 5.3. Psycholinguistics
- 5.4. Corpuslinguistics

0.1. Grammar = linguistics from school:

Concepts and categories of traditional grammar (from Aristotle to Quirk)

0.1.1. Classification parts of speech (POS = word classes)

major - minor:

- productivity
- diachronic change
- semantic: +/- lexical

VERBS
NOUNS
ADJ.
ADVERBS
PREP
CONJ
PRON
AUX / NUM / ART
(INTERJECTIONS)

0.1.2. Establishing patterns/paradigms (secondary categories)

N: declension
V: conjugation

GENDER (# SEX)	
CASE	N
NUMBER	
PERSON	
TENSE (# TIME)	V
MOOD	
VOICE (= GENUS VERBI)	

0.1.3. Functional elements (= clause elements) in a particular sentence

ADVERBIAL	Tomorrow	To everyone's surprise
SUBJ	Mrs. Thatcher	she
VERB/PREDICATE	will introduce	was elected
OBJ dir/indir	her cabinet	
	to the Queen	
COMPLEMENT		Prime Minister
subj/obj		
		reference identity

0.2. Linguistics = thinking about language

0.2.1. Linguistics

0.2.1.1. History of linguistics (schools)

- traditional grammar
- neogrammarians
- structuralism
- generative-transformational grammar
- cognitive grammar

0.2.1.2. Subdisciplines

cf. the football model

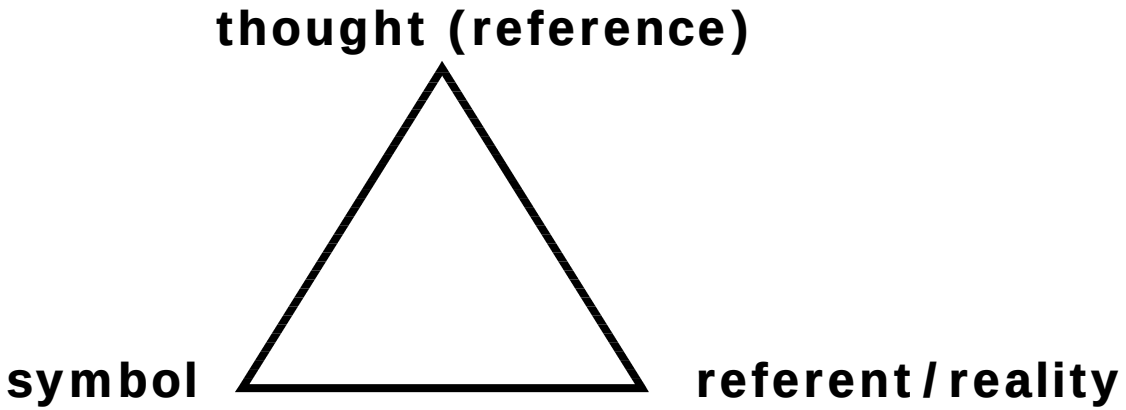
- microlinguistics: phonology, morphology, lexicography, semantics, syntax, text analysis
- macrolinguistics: pragmatics, sociolinguistics, psycholinguistics, etc.
- applied linguistics: lexicography, translation studies, error analysis, computer linguistics

0.2.2. Language

0.2.2.1. Definitions of language

Language is a purely human and non-instinctive method of communicating ideas, emotions and desires by means of voluntarily produced symbols (Sapir 1921)

language and the linguistic symbol/semiotic triangle (de Saussure):



Communication vs. language

communication = *the passing on or exchange of information – distinguishes what is living from what is non-living in nature* (O'Grady et al. 1996)

human language and animal communication:

the **design features of human language**

1. interchangeability: all members of the species can send and receive messages
2. feedback: users of the system are aware of what they are transmitting
3. specialization: the communicative system serves no other function but to communicate

The design features of human language 2

4. semanticity: the system conveys meaning through a set of fixed relationships among signifiers, referents and meaning.
5. arbitrariness: there is no natural or inherent connection between a token and its referent
6. discreteness: the communication system consists of isolatable, repeatable units
7. displacement: users of the system are able to refer to events remote in space and tine
8. productivity: new messages on any topic can be produced at any time
9. tradition, cultural transmission: certain aspects of the system must be transmitted from an experienced user to a learner

The design features of human language 3

9. duality of patterning: meaningless units (phonemes) are combined to form arbitrary signs. signs can be recombined to form new larger meaningful units (s-p-o-t **à** tops, pots)
11. prevarication: the system enables users to talk nonsense or to lie
12. learnability: the user of the system can learn other variants. Humans can learn different languages, bees are limited to their genetically specified dialect
13. reflexiveness: the ability to use the communication system to discuss the system itself

0.2.2.2. Types of languages

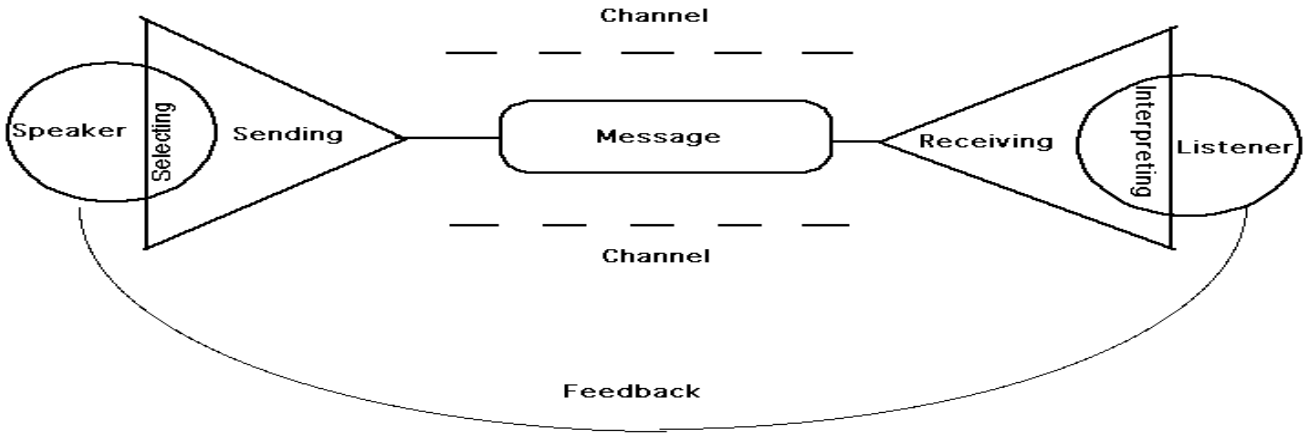
relation between words expressed

- in different words = analytic/isolating (Chinese)
- in merged/fused affixes = synthetic/inflecting (Latin)
- in unchanging affixes = agglutinative (Turkish)
- others, like polysythetic (Inuit, Nutka)

diachronic shifts: English from synthetic to analytic (further than German)

0.2.2.3. An ideal model of communication

- channel: the messages are primarily transmitted via the vocal-auditory channel
- linearity: the message is extended temporally (speech) and locally as a string (writing) and is produced and analyzed as a sequence.
- redundancy: the same information given several times



0.2.2.4. The wild world of Englishes

ENL = English as a native language

ESL = English as a second language

EIL = English as an international language

lingua franca (*lingue franche*) of international science and technology (ESP = English for specific purposes)

EAP = English for academic purposes

pidgin languages develop rudimentary grammar in superficial contact situations (Tok Pisin)

creole languages develop from pidgins when they expand in form, parallel to more functions as a first language (Krio)

1. Phonetics and Phonology

1.1. Introduction to phonetics

Phonetics = t*he study of the speech sounds that occur in all human languages to represent meanings.*
(Fromkin/Rodman 1993:176)

types of phonetics

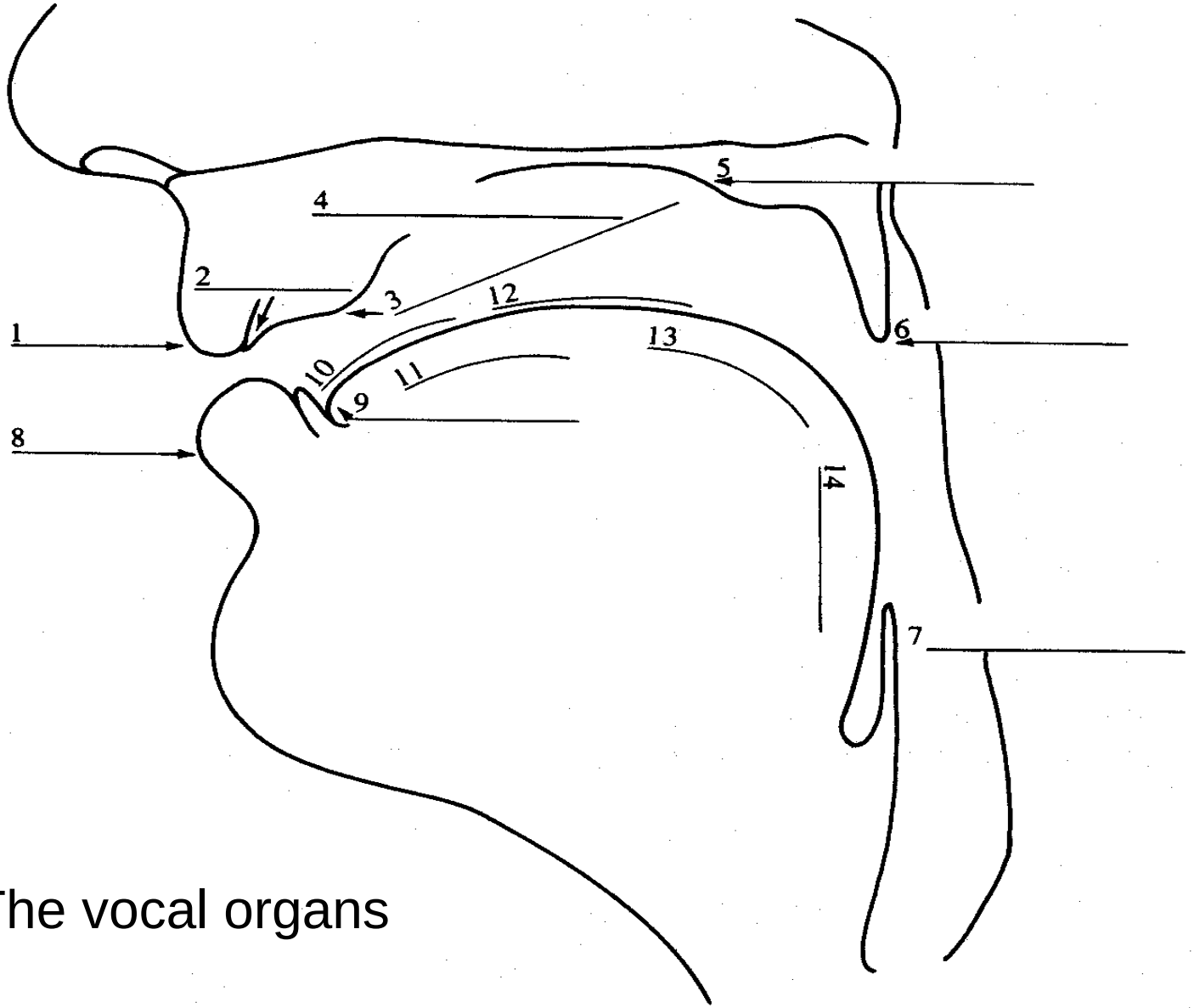
1. articulatory phonetics
   study of the way how speech sounds are made (articulated) by the vocal organs
2. acoustic phonetics
   study of the physical properties of the speech sounds
3. auditory phonetics
   study of the perceptual response to speech sounds through ear, auditory nerve, brain

1.2. Vocal organs and articulators

area above larynx: vocal tract
parts of the oral tract forming sound:
articulators
front to back: lips, teeth, alveolar ridge,
hard palate, soft palate (velum)
velum: a flap that can shut off the nasal tract
end of velum: uvula

part between larynx and uvula: pharynx

tongue: can be separated into: tip, blade, front, center, back



The vocal organs

1.1.3. Articulation of consonants and vowels

vocal cords: can vibrate under pressure of airstream
vibrating cords = voiced
non-vibrating cords = voiceless

• vowels vs. consonants:
vowels: little obstruction of airstream, generally voiced
= continuous “Selbstlaut”
consonants: voiceless or voiced, obstructed airstream
-- > consonants to be classified according to place and manner of obstruction

1.1.3.1. Consonant articulation

Places:

1. labial/bilabial (upper and lower lips)
   <pie>, buy>, <my>
2. labiodental (lower lip + upper front teeth)
   <fire>, <fun>, <vicious>
3. dental/interdental (tongue tip + upper front teeth)
   <thigh>, <thy>
4. alveolar (tongue tip/blade + alveolar ridge)
   <tie>, <die>, <lie>
5. retroflex (tongue tip + back of alveolar ridge)
   <rye>, <row>, <ray> and <hour>, <air>

• not used by all speakers of English

Places of consonant articulation 2

6. palato-alveolar (tongue blade + back of alveolar ridge)
   <shy>, <she>, <show>
7. palatal (tongue front + hard palate)
   <Hugh>
8. velar (tongue back + soft palate)
   <hack>, <hag>, <hang>
9. glottal (vocal cords)
   <heave>, <hug>

not used in English:
uvular (French <r>); pharyngeal (Arabic); clicks (Zulu)

Manners:

articulators can close the oral tract completely or partially

1. stop (closure, airstream cannot escape)

• nasal stop: air stopped in mouth but can escape through nasal tract <my>, <night>, <song>

oral stop: raised velum closes nasal tract **à** pressure builds, airstream is released in bursts:

<pie>, <cool>, <guy>, <tool>

2. fricative (close approximation of two articulators)

airstream is partially obstructed **à** turbulent airflow **à** hissing sounds

<shy>, <those>, <friend>

• higher-pitched: sibilants

• lower-pitched: non-sibilants

Manners of consonant articulation 2

3. approximant

• narrowing of articulators until turbulent airstream occurs but not close enough for a fricative

<we>, <Howard>

4. lateral

obstruction along center of oral tract without complete closure <lip>

5. affricates

some sounds are combinations of other simpler sounds, cf. <church>

stop + fricative = affricates

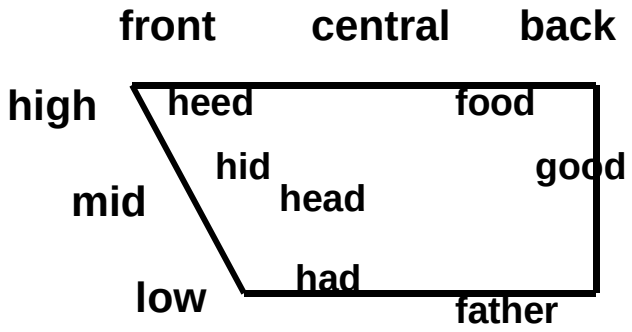
1.1.3.2. Articulation of vowels

• articulators are open, airstream unobstructed
cf. <heed, hid, head, had, father, good, food>
• tongue tip on front lower teeth
dome of tongue: raised
<heed, hid, head, had>: highest point of tongue: front of mouth **à** front vowels
high front vowels <heed> and low front vowels <had>
• mouth is increasingly open

tongue close to back of vocal tract **à** back vowels
high back vowels <food> and low back vowels <father>

Articulation of vowels 2

lip position: close together in mid and high back vowels <good, food>
lip rounding: rounded vs. unrounded vowels
à three factors for vowels
1. height of the body of the tongue
2. front-back position of tongue
3. degree of lip rounding



relative position of the highest point of the tongue

1.1.3.3. Articulation of diphthongs/triphthongs

= a glide from one vowel position to another
(less extreme)
in E: all diphthongs are centreing or falling

triphthongs (centering)
are unstable and tend to be reduced to monophthongs:
shire, *shower* > [sha:] *homophones!*

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)

CONSONANTS (PULMONIC)

	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				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.

CONSONANTS (NON-PULMONIC)

Clicks		Voiced implosives		Ejectives	
ʘ	Bilabial	ɓ	Bilabial	'	as in:
|	Dental	ɗ	Dental/alveolar	pʼ	Bilabial
!	(Post)alveolar	ʄ	Palatal	tʼ	Dental/alveolar
ǂ	Palatoalveolar	ɠ	Velar	kʼ	Velar
ǁ	Alveolar lateral	ʛ	Uvular	sʼ	Alveolar fricative

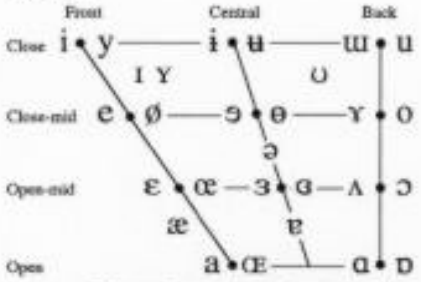
SUPRASEGMENTALS

ˈ	Primary stress	ˌfoʊnəˈtɪʃən	
ˌ	Secondary stress		
ː	Long	eː	
ˑ	Half-long	eˑ	
̆	Extra-short	ĕ	
.	Syllable break	ɹi.ækt	
|	Minor (foot) group		
‖	Major (intonation) group		
‿	Linking (absence of a break)		

TONES & WORD ACCENTS

LEVEL		CONTOUR	
é or ˥	Extra high	ě or ˩˥	Rising
é ˦	High	ê ˥˩	Falling
ē ˧	Mid	ĕ ˧˥	High rising
è ˨	Low	ě ˨˩	Low rising
ȅ ˩	Extra low	ȇ ˥˧	Rising-falling
↓	Downstep	↗	Global rise
↑	Upstep	↘	Global fall

VOWELS



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

OTHER SYMBOLS

ʍ	Voiceless labial-velar fricative	ɕ ʑ	Alveolo-palatal fricatives
w	Voiced labial-velar approximant	ɺ	Alveolar lateral flap
ɥ	Voiced labial-palatal approximant	ʧ	Simultaneous ʃ and x
ʜ	Voiceless epiglottal fricative		
ʢ	Voiced epiglottal fricative		
ʡ	Epiglottal plosive		

Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary.

k͡p t͡s

DIACRITICS

Diacritics may be placed above a symbol with a descender, e.g. ŋ̊

̥	Voiceless	n̥ d̥	̤	Breathy voiced	b̤ a̤	̪	Dental	t̪ d̪
̬	Voiced	s̬ t̬	̰	Creaky voiced	b̰ a̰	̺	Apical	t̺ d̺
ʰ	Aspirated	tʰ dʰ	̼	Linguolabial	t̼ d̼	̻	Laminal	t̻ d̻
̹	More rounded	ɔ̹	ʷ	Labialized	tʷ dʷ	̃	Nasalized	ẽ
̜	Less rounded	ɔ̜	ʲ	Palatalized	tʲ dʲ	ⁿ	Nasal release	dⁿ
̟	Advanced	u̟	ˠ	Velarized	tˠ dˠ	ˡ	Lateral release	dˡ
̠	Retracted	i̠	ˤ	Pharyngealized	tˤ dˤ	̚	No audible release	d̚
̈	Centralized	ë	~	Velarized or pharyngealized	ɫ			
̽	Mid-centralized	e̽	̝	Raised	ę (ɹ̝ = voiced alveolar fricative)			
̩	Syllabic	n̩	̞	Lowered	ę (β̞ = voiced bilabial approximant)			
̯	Non-syllabic	e̯	̘	Advanced Tongue Root	e̘			
˞	Rhoticity	ɔ˞	̙	Retracted Tongue Root	e̙			

1.2. Phonology

1.2.1. Introduction

speech sounds to be analyzed after:

physical properties (form) **à** phonetics

sound differences / similarities (function) **à** phonology

phonetics	**phonology**
sounds of language	functioning of sounds as part of a system
parole, speech act	langue, language system
universal	language specific
concrete	abstract
phone [ ]	phoneme / /

• sounds form segments; speakers know which segments contrast, i.e. are in opposition or distinctive

sip vs. *zip*; *hit* vs. *hot* **à** minimal pairs

= 2 forms with distinct meanings that differ only by one segment

1.2.2. Levels of description: from minimal pairs to phonemes

- established on basis of sound, not spelling
- only one segment can differ, NOT *soldier* vs. *shoulder*
- contrasts are language-specific, i.e. sounds that are distinctive in one language may not be distinctive in another

wide vs. narrow transcription for *leaf-feel*
[l] is never to differentiate meanings (cf. 1.2.3. below) **à** difference is phonetic, not phonemic:

- unit of description: phoneme /l/

phoneme: smallest unit with potentially distinctive function
variants: allophones, cf. German /x/: *ich* vs. *Buch*

1.2.3. Principles in phonology

• complementary distribution: phonetic units that never occur in the same environment = allophones if phonetic similarity
[l] only in front of vowels and /y/: clear
[l] in front of consonants and word endings: dark
• free variation: <economics> phonetic difference realised by speakers for the same word

spelling systems generally ignore phonetic variation that is non-distinctive, evidence that speakers have a mental notion what phonemes are
phonologically relevant differences are never left out in spelling:
cf. /r/ and /l/ in *rift* vs. *lift*

• neutralization: foreigners can have difficulty in phonological difference, cf. German *Auslautverhärtung*
German: *Rad* vs. *Rat*

1.2.4. Phoneme relationships

• linking (liaison): BrE (is non-rhotic, but SW, Shakespeare!) avoids two distinct vowel phonemes

à insertion of liquid [r] or glide [j] or [w]

near – nearing	*near Africa*
see – seeing	*to see Arthur*
sue – suing	*to sue Arthur*

intrusive /r/ in *law and order*

• phoneme relationships:

/-et/	/p-t/	/pe-/
/p/	/e/	/t/
/b/	/i/	/n/
/l/	/o/	/k/

= matrix of real and potential words

à language can contain irregular words: as loan words, foreign words

1.2.5. Distinctive features of English stops

	/k/	/g/	/ŋ/	/p/	/b/	/m/	/t/	/d/	/n/
La	-	-	-	+	+	+	-	-	-
Ve	+	+	+	-	-	-	-	-	-
De	-	-	-	-	-	-	+	+	+
St	-	+	-	-	+	-	-	+	-
Na	-	-	+	-	-	+	-	-	+

sub-phonemic analysis
basis: distinctivity of the 9 phonemes
phonemes of one language: can only be defined in contrast to other phonemes of the same language (Hockett)

1.3. Suprasegmental phonology

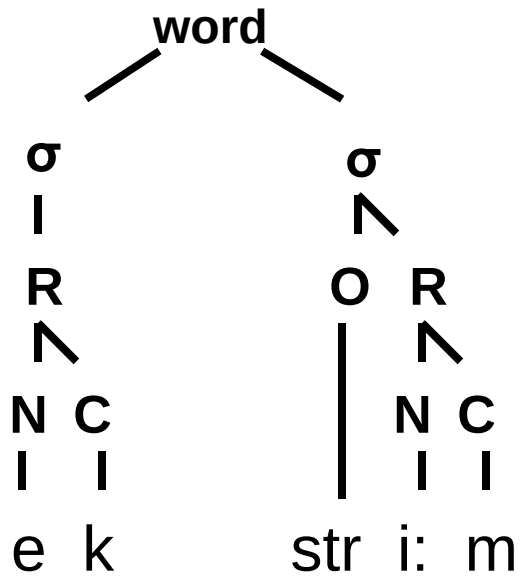
1.3.1. Levels of description: syllable

syllable = composed of a nucleus (usually a vowel) and its associated non-syllabic elements
nucleus (N): syllable's obligatory member, forms core
coda (C): consists of those elements following the nucleus in the same syllable
rhyme (R): nucleus + coda
onset (O): elements preceding the rhyme
reason: speakers syllabify after underlying rules
• closed vs. open syllable: syllable with coda vs. syllable without coda

Syllabification of words

1) identify nucleus: obligatory, each vowel makes a syllabic nucleus
2) longest sequence of consonants to the left that does not violate phonotactic rules: onset
3) remaining consonants to the right: coda

extreme:



1.3.2. Phonotactics

= a set of constraints on sound combinations

how segments are formed is part of speaker's knowledge of his/her language

similar to German:
word-initial consonant clusters /str-/, /spr-/, /sl-/, /sm-/, */sfr-/ (difficult for other foreigners)

but also differences from German, because
- sound changes (*knight, write*)
- foreign words are accepted or adjusted (*psychology*)

1.3.3. Introduction to prosody

word stress: BrE *ˈse.kre.tri* - AmE *ˈse.kreˈtari*

rhythm: isochrony / English is **stress-timed**
= same time span between **stressed** syllables
contrast syllable-timed = same between all syllables
(French, African languages)
à weak forms in unstressed position:
auxiliaries, prepositions/conjuncts, pronouns/determiners
features of **connected speech**:

- function words: *he's* vs. *he is*; *he'll* vs. *he will*
- **assimilation** = adjacent sounds are altered in cotext to make them more similar, i.e. easier to pronounce

types: partial [*tem baiks*] – total [*tem mais*]
regressive [*speiship*] – progressive (rare!) – coalescent [*wudshu*]

1.3.4. Intonation

= variation and control of pitch has 3 functions:

grammatical = to distinguish declarative (falling) from interrogative clauses (rising)

pragmatic = to manage information, emphasizing NEW vs. old information

attitudinal = to signal emotions (surprise/enthusiasm=rise-fall, uncertainty/doubt=fall-rise, boredom/irony/sarcasm=level tone)

cf. "*great*"

English is not a tone language like Chinese

(*ma* may mean "mother" or "horse")

1.3.5. Introduction to orthography

English has fewer graphemes than phonemes >
digraphs <sh>, <ch>, <th>, double graphs for length */door/*
huge discrepancy between writing and pronunciation:

- 1 phoneme many graphemes:
  /o:/ in *glory*, *all*, *door*, *bought*,
- 1 grapheme many phonemes:
  <ou> in *stout*, *soul*, *bought*, *draught*, *youth*, *young*

in English for historical reasons:
orthography fixed earlier than German (Caxton 1476)

- silent letters: <gh>, final <-e>
- different conventions (e.g. <VCe#> = long vowel in *mate/mat*)
- Great Vowel Shift (continuous / 1500 - 1700?)
  = long vowels become closer and diphthongised

-> spelling reform?

2. Morphology

2.1. Introduction

Morphology deals with the internal structure of words that can be broken down into meaningful parts

à concerned with how speakers understand and create complex words

words have internal structure consisting of smaller units

morpheme = smallest unit that carries meaning / information about function

= smallest unit in grammatical analysis

word is unreliable in English because of unclear spelling rules: *football* vs. *gold watch* (NOT *golden watch*)

pronunciation decides: *'black 'bird* vs. *,black 'bird*

NB: languages differ in morphological complexity:

Japanese low, Turkish high

2.2. Morpheme types

2.2.1. Allomorphs

= a group of morphs constituting 1 morpheme

variation is phonologically or lexically conditioned

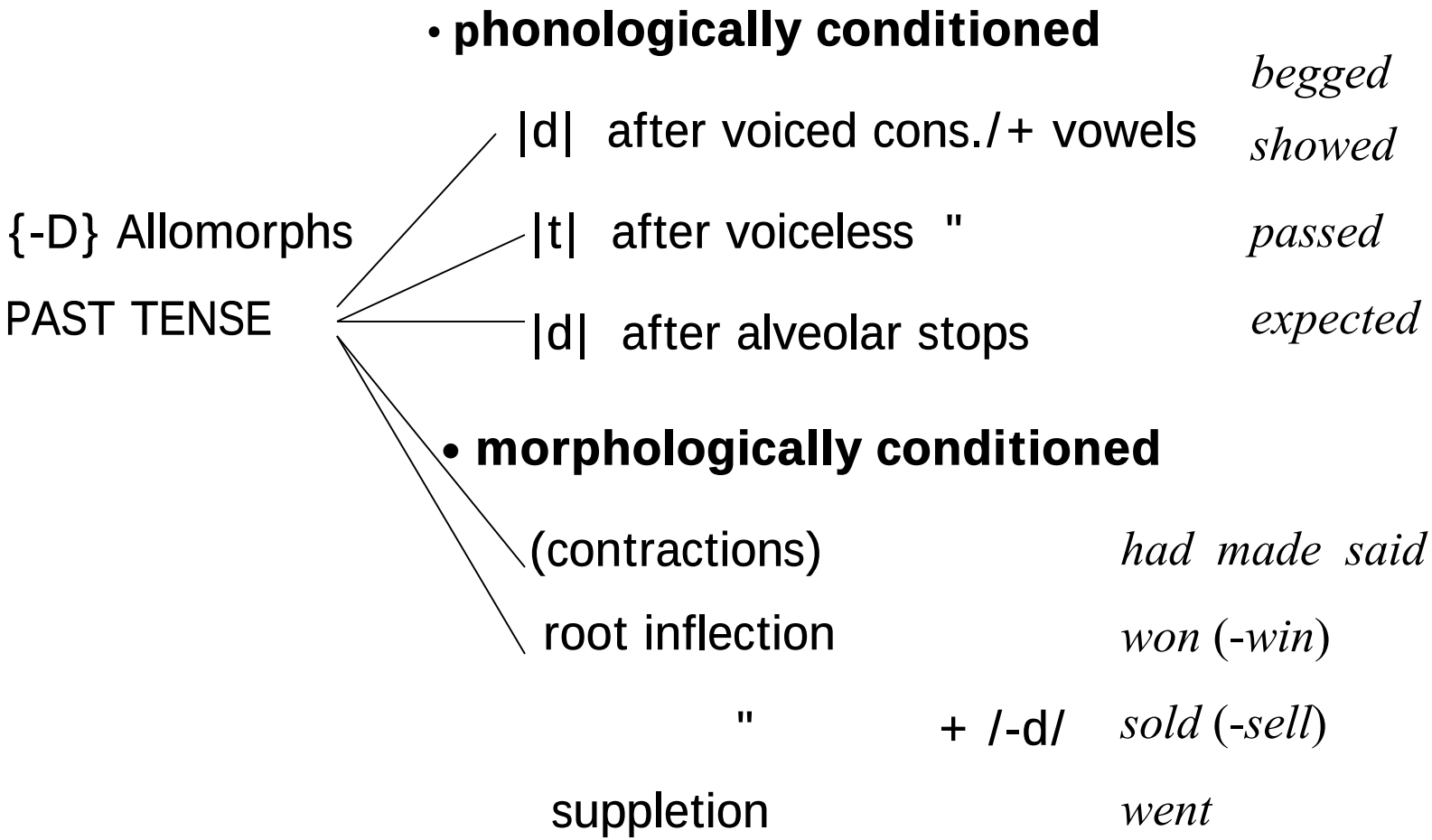
e.g. plural {-*S*}
past tense {-*D*}
past passive participle {-*N*}

special cases:

zero morpheme = no form, but meaning (*sheep-sheep)*

empty morpheme = form, but no meaning (*to do*)

discontinuous morpheme = 2 forms, 1 meaning (*is –ing*)



2.2.2. Form - function relationships

build-er; *marry/remarry*: phonology/form is not revealing

- *–er* +NOUN: indicates function of word: agent cf.
  reader, writer, runner
- *re-* +VERB: meaning “again” understood automatically cf.
  reconsider, rebuild

but restrictions: **relike*, **rehave*

immediate constituants:
|*criminal law*|²*yer*|³, |*heavy smok*|²*er*|³

2.2.3. Morpheme level of analysis

simple/monomorphemic words: no further subdivision
complex/polymorphemic words: 2 or more morphemes

• basic types: free vs. bound morphemes
free = a morpheme can be a word by itself
doghouse, ready-made
bound = a morpheme must be attached to another element
{{*un*{{*manage*}$_V$ *able*}$_A$}$_A$}*ness*}$_N$
• lexical vs. grammatical morphemes
lexical (in lexicon) for constructing new words {*black*}{*bird*}
grammatical (in grammar)
bound + lexical = derivational (*disbelief, readable*...)
possible: change of word class, change in meaning

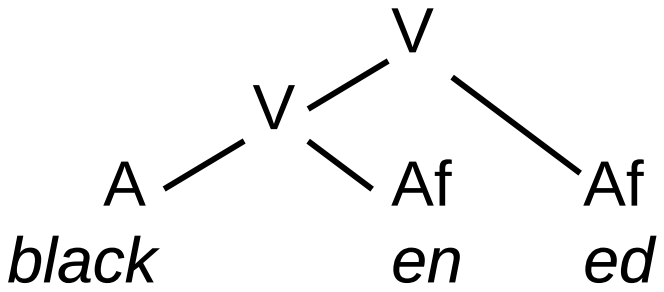
2.3. Inflectional morphology

= what forms a word can take depends on role in sentence
grammatical morphemes express grammatical relationship between word and context: plural-*s*, past tense *–ed*
free grammatical morphemes = function words *and, the*

Inflectional morphemes in English	
• nouns:	
plural *–s*	*the books*
• verbs:	
3rd person sg. non-past *-s*	*Mary reads well*
progressive *–ing*	*John is working*
past tense *–ed*	*She read*
past participle *–en/ed*	*He has eaten/worked*
• adjectives:	
comparative *–er*	*taller*
superlative *–est*	*tallest*

2.4. Morphological structure of words

à necessary to identify and classify morphemes according to function for the word and its meaning
complex words: root + one or more affixes
• root morpheme: major component of word's meaning, usually root belongs to N, V, A, P
affixes: always bound morphemes
• base: the form to which a morpheme is added



A: root and base for -*en*; V: base for –*ed*
example/exercise: *unhealthy, pretreatment*

2.4. Word formation

2.4.1. derivation

= process of morphological variation in the constitution of words

morphological patterns vary in degree of **productivity**

high = many: *-ness* (*brightness*) vs. low = few *-ity* (*eternity*)

unproductive: *-dom* (*kingdom, martyrdom,* etc.)

2.4.2. composition

= combination of 2+ lexical morphemes *black+bird* **à**

different lexical categories combinable

endocentric composition: right morpheme determines word class: *blackbird, spoonfeed, nationwide*

exocentric: meaning cannot be inferred from rightmost component (*walkman, redneck*)

spelling difference: *oak leaves vs. Toronto Maple Leafs*

2.4.3. zero-derivation / conversion

= change of word class without change in form

challenge – to challenge, ship – to ship

V derived from N, N derived from V

less common: N from A (*the poor*)

V from Prep (*to down a beer*)

2.4.4. clipping and blending/telescoping

= shortening of polysyllabic words

laboratory > lab, gymnasium > gym; influenca > flu

smoke + fog = smog, breakfast + lunch = brunch

2.4.5. backformation

= to remove a real or potential affix

to housekeep, to babysit

2.4.6. acronyms

= initial letters of words combined (in capitals)

(first isolated in pronunciation, later combined if possible)

NATO, LASER, AIDS

3. Syntax
3.1. Introduction

syntax = how people combine words to form sentences

- speakers: finite set of memorized words/ morphemes
  as basis for potentially infinite sets of sentences
- *discrete infinity* (recursive)
  = basis of creativity of human language
- allows speakers to create/understand novel sentences

syntactic theory = how speakers know how to form
sentences and how they get this knowledge

- speakers' knowledge: mental grammar

3.1.1. Syntactic structure: constituents and word order

• language: structured, not random **à** rules = grammar
concerns of syntax: word order
• words behaving as units: constituents
The cat ate the rat / The rat ate the cat.
à same words, different meaning
• speakers "know" about importance of word order
but: *the rat, the cat ate* = larger units than words
groups of words forming a unit: in [...]

[*our vicar*] – [*likes*] – [*fast cars*] – units, because:
[*he*] - [*likes*] – [*them*]
[*our vicar*] - [[*likes*] – [*fast cars*]]
a unit because *like* is a V_t

3.1.2. Forms and functions

function of constituents: difference in meaning

The cat ate the rat/ The rat ate the cat

• subject: performs action, is agent, what the sentence is about

• predicate: what subject is engaged in doing, predicate is anything except subject

à operations of finding subjects: simple, formal subjects not always "do" something (agentive in case grammar)

I dislike the idea. Miriam stood aside.

can be meaningless: *It was hot, It is raining*

There are ways of making you talk.

à non-referential it and existential there: fill subject slot

3.2. Functions

3.2.1. Properties of subjects vs. objects

subjects = predominantly nouns, groups with N: NPs
(*the stupid dog, the girl with the red hair, this committee...*)
subjects are
a) usually NPs
b) (usually) the 1st NP in the clause
c) obligatory
d) determine forms of verbs (agreement in 3rd PSingPres)

objects are
a) often NPs
b) after V

direct object (DO) = entities that undergo process denoted by verb:
He broke the teapot.
• play patient role (= semantic test)

complements denote the same referent as subject or object

3.2.2. Functions of direct vs. indirect object

DO (active sentence) **à** subject (passive sentence)
DOs complete the meaning of the verb, are complements
à complement: any element that is required by another element
• **indirect objects** (IO): typical role: receiver, goal
We gave the boys the CDs.

verbs taking DO and IO: ditransitive verbs
a) usually NPs
b) cannot occur without DO
c) always precede DO in E (not in German)
d) can be passive subjects (*The boys were given the CDs.*)

3.3. Form

3.3.1. Word forms, word classes, phrases

words: difficult to define: *dogs, eats, duty-free*
à grouping into word classes, parts of speech

• word classes are notions of form, not function
• criteria of nouns; words preceded by: *a ,the, this...*
common determiners: *the/a, this/these, that/those,*
• a noun can be preceded by adjectives
à N are characterized by their environment
subclasses: common N (+/-count), proper N, numerals (cardinals/ordinals), pronouns
• nouns are heads of NP: *the hat, blue hat on the shelf*
hat = central element (Head)

3.3.3. Form criteria of verbs and adjectives

verbs

inflections encode grammatical properties (*-ed* **à** past) like tense, agreement

• main verbs and auxiliaries:

aux. express point of view (deontic vs. epistemic)

non-finite verbs: *to*-infinitive (*I wanted him to dance),*

participles (wanting, reconsidered)

V are head of VPs: *The library* [$_{VP}$ *recalled their books*]

adjectives

formal markers: *-ful, -ible, -ive* but not exclusive: *green*

a) are gradable (*very*...) – exceptions: materials, Nationalities (*?very wooden, ?very Swedish*)

b) can take comparative, superlative

exceptions: *good-better-best* = analytical comparison

predicative: with V_L: *appear, be, feel, look, seem, smell*

A are Head of APs: [$_{AP}$ *very glad to be here*]

3.3.4. Form criteria of prepositions, adverbs and conjunctions

adverbs: modify verbs, adjectives or other adverbs; or clauses!
-ly, -wards, -wise, -ways, but not all (*very, here, now*)
some have comparison (*well, soon*)
classes: circumstantial *often, reluctantly*
degree *extremely, very*
sentence *however, probably, perhaps*

prepositions have no formal criteria
are head of PPs [$_{PP}$ *with* [$_{NP}$ *the dog*]]
often: NPs as prepositional object/prepositional complement

conjunctions have a linking function
a) coordinating: *and, or, but*
b) subordinating: *that, if, whether, for, because*

3.4. Clauses and sentences

clause: a self-containing expression which contains a subject and a predicate

main vs. subordinate / matrix vs. subclause **à** complete sentence

finite vs. infinite clauses:

+/- *to* infinitive,

present/past participle *–ing/-ed*

most cases: predicate has a finite lexical verb

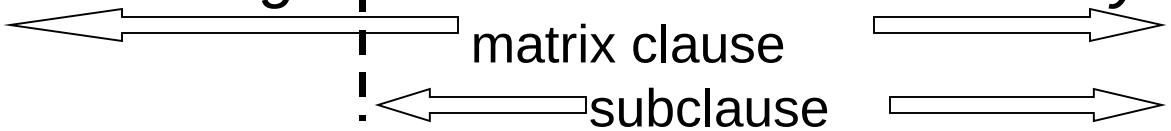
à number of lexical verbs ≅ number of clauses

a) *I paid the entire bill at once*.

b) *They were happy after I had paid the bill at once*.

c) *They wanted me to pay the entire bill at once*.

Tim thought ¦ *that Kate believed the story*.



that: complementiser

3.5. Theoretical approaches to syntax

3.5.1. Structuralist grammar

de Saussure --> American Indian anthropology:
Boas - Sapir - Whorf - Bloomfield (*Language* 1933)

Sapir - Whorf hypothesis (cf. W. von Humboldt):
linguistic determinism + relativity =
language determines/influences culture

emphasis on methodology (corpus-based)

segmentation - classification:

- immediate constituent (IC) analysis (e.g. *heavy smoker*)
- bracketing and tree diagrams

application (in language lab): pattern practice, sentence switch boards

3.5.2. Transformational grammar

Noam Chomsky: 1957 *Syntactic Structures*
1965 *Aspects of the Theory of Syntax*

language = an infinite number of sentences, generated by applying an unconscious finite system of rules (rule-governed creativity)

ideal speaker-listener: performance - competence
(all potential well-formed sentences)

selection restrictions (e.g. **colourless green ideas sleep furiously*)

generate (in the mathematical sense) = rewrite rules x->y (rewrite x as y):
e.g. S --> NP+VP; VP --> V+NP

transformations: surface structure --> deep structure
e.g. passive transformation: NP1 + VP + NP2 --> NP2 + *be* VP (+ *by* NP2)

solves ambiguities (e.g. *murdering peasants can be dangerous*)
agens/agentive or patiens/object?

3.5.3. Case/valency grammar

underlying cases (e.g. agentive, instrumental, dative/benefactive, objective, locative, source, goal) matched by surface-structure relations

verb as central element with obligatory complements: e.g. *NP put NP PNP*

obligatory complements, e.g. *They* enjoyed *the trip*.

optional complements, e.g. She is reading *a book*.

adjuncts, e.g. Alcohol can be obtained *at the bar*.

1. John opened the door with the key.
2. The door was opened by John.
3. The key opened the door.d
4. The wind smashed the window.
5. John gave me the book.
6. I went to London the other day.
7. This tent sleeps ten people.
8. No one can escape the difficulties of life.

3.5.4. Cognitive grammar

"Cognitive linguistics is an approach to the analysis of natural language that focuses on language as an instrument for organizing, processing, and conveying information. (...) The formal structures of language are studied not as if they were autonomous, but as reflections of general conceptual organization, categorization principles, processing mechanisms, and experiential and environmental influences." (Geeraerts 1995: 111)

conceptualisation in grammar

cf. *gold nugget* vs. *gold dust* – continuum of sizes

BUT grammar does not reflect physical reality but experienced reality

à conceptual distinction:

grammar codes *gold nugget* as a count noun, bounded

grammar codes *gold dust* as a mass noun, unbounded

conceptualisation in lexicon: metaphor (cf. Lakoff)

Lakoff, George/Mark Johnson (1980), *Metaphors we live by*. Chicago: U. of Chicago Press. pp. 3-24

1 argument is war
2 time is money
2a time is a limited resource
2b time is a valuable commodity
3 ideas (or meanings) are objects
4 linguistic expressions are containers
5 communication is sending
6 happy is up; sad is down
6a conscious is up; unconscious is down
6b health and life are up; sickness and death are down
6c having control or force is up; being subject to control or force is down
6' more is up; less is down
6d foreseeable future events are up (and ahead)
6e high status is up; low status is down
6" good is up; bad is down
6f virtue is up; depravity is down
6g rational is up; emotional is down
6h active is up; passive is down
7 the mind is an entity
7a the mind is a machine
7b the mind is a brittle object

4. Semantics

4.1. Introduction

4.1.0. Ogden/Richards (1923). *The meanings of meaning*

4.1.1. For language to fulfill communicative function/convey a message:
form must have content

same form – different content: ambiguous sentences:
cf. *Ruth saw the people with binoculars.*

4.1.2. structuralist system:

seme --> sememe + morpheme = lexeme?

4.1.3. approaches:

semasiological = FORM ---> CONTENT
e.g. *chair* means 1. "thing to sit on" 2. professor

onomasiological = CONTENT ---> FORM
e.g. "things to sit on" are called: *chair*, *arm-chair*, *stool*, *sofa*, *couch*, etc.

4.1.4. 7 types of meaning

example

1. CONCEPTUAL MEANING (or *Sense*)		Logical, cognitive, or denotative content	(cf. theories)
ASSOCIATIVE MEANING	2. CONNOTÀTIVE MEANING	What is communicated by virtue of what language refers to.	female= soft, caring, likely to cry, cooking?
	3. SOCIAL/ STYLISTIC MEANING	What is communicated of the social circumstances of language use.	domicile – residence – abode – home; cast – throw - chuck
	4. AFFECTIVE MEANING	What is communicated of the feelings and attitudes of the speaker/writer.	will you belt up [+intonation]
	5. REFLECTED MEANING	What is communicated through association with another sense of the same expression.	intercourse, erection, gay
	6. COLLOCATIVE MEANING	What is communicated through association with words which tend to occur in the environment of another word.	heavy smoker, rain, fine; pretty vs. handsome
7. THEMATIC MEANING		What is communicated by the way in which the message is organized in terms of order and emphasis.	She donated the first prize The first prize was donated by her

Source: Leech, Goeffrey (1981). *Semantics.* 2nd ed. Harmondsworth: Penguin, p.23.

4.2. Semantic theories

4.2.1. Semantic features/markers theory (Fodor/Katz)

list components (incl. distinctive markers):

e.g. *cat* [+concrete +animate -human +mammal +mature +/-male]

the meaning of single words is determined in componential analysis

- *girl* [+anim, +human, -adult, +female]
- *woman* [+anim, +human, +adult, +female]
- *table* [-anim]

	aunt	girl	calf	. . .
animate	+	+	+	
human	+	+	-	
male	-	-	+-	
mature	+	-	-	
⋮				

field theory (Trier 1930)

= field value due to oppositions to paradigmatic neighbours

- *walk*, *march*, *pace*, *amble*, *stroll*, *sneak*, *stagger*
- *stupid*, *thick*, *silly*, *dumb*

mosaic?? overlaps, lexical gaps

e.g. English: Shona or Welsh colour terms (cf. 4.2.2)

ENGLISH	green	blue	grey	brown
WELSH	gwyrdd	glaz		llwydd

4.2.2. Prototype theory: the emergence of prototypes

- borders of meanings: blurred, fuzzy

bird [+animate, -human, +wings?, +lays eggs, +can fly?, +feathers?]

à concept of prototypes

use of attributes: can be similar (*birds*) or not (*games*)

board/ball/card games only a network of overlapping similarities

"family resemblance" (cf. Wittgenstein)

- attribute tests confirm the (intuitive) "best example"
- thus: prototypical members have largest number of attributes in common

example: basic color terms (defined by brightness, hue, saturation); focal colors are consistent for speakers of the same and of other languages (concentric circles: *white*, *black*, *red*; *blue/green*, *yellow*; *brown* ...)

(4.2.2) prototype = “the clearest cases of category membership defined [...] by people’s judgements of goodness of membership in the category” (Rosch)

- humans classify things into categories with no discrete boundaries
- categories can be distinguished with emphasis on their structure
- prototype: an image that averages similar experiences

most frequent phenomena: coded as basic categories
earliest to be learned / easiest to be triggered
classic example: *bird* prototype: *robin*

4.3. Semantic relations

4.3.1. Synonymy and antonymy

synonymy = two words have the same meaning in a number of contexts: *I spent my holidays/vacations in Spain* but *Christmas, Easter holidays*

real synonymy: rare or not-existent

youth	-	adolescent
purchase	-	buy
remember	-	recall
begin	-	start

antonymy = lexemes contrast in semantic feature(s)

- one member can be marked: *How tall is Rita?*
  (*tall* vs. *small*, *tall* is unmarked)

graded antonymy: *not clever ≠ stupid*

ungraded antonymy: *alive* vs. *dead*

dark	-	light
hot	-	cold
in	-	out

4.3.2. Polysemy and homonymy

polysemy = lexemes can have two or more related meanings cf. *surfer*

à to be seen as single word with different meanings

bright:	shining	– intelligent
deposit:	minerals in the earth	– money in the bank

homonymy = lexemes have entirely distinct meanings

à to be seen as separate words with same pronunciations

bat:	flying mammal	– equipment in baseball
club:	social organization	– a blunt weapon

homography = words are written identically but pronounced differently: *wind*

homophony = words are pronounced identically but written differently: *threw – through*

4.3.3. Syntagmatic Relations

collocations = words which tend to occur together: *fair hair, fair play*
collocational range = a lexeme has few collocates
(*omen* =>only: *good*, *bad*; *good* has many)
idioms = composite meaning cannot be deduced from individual elements (semantically opaque): *to pull s.o.'s leg*
sayings: *the early bird catches the worm*
rituals: *keep one´s fingers crossed*

4.3.4. Non-lexical semantics

meaning in grammar: continuous form

- aspectual meanings:
  - progressive: *he was reading*
  - imperfective: *she was phoning*
  - incomplete: *he has been studying*
  - duration iterative: *they were nodding*
  - dynamizing (stative>dynamic): *he is being silly*
    - temporary: *he is cycling to work* (*his car is in the garage*)
    - vaxing/vaning: *you are making more and more/fewer and fewer mistakes*
- time meanings: near future plan: *I'm phoning her tonight*

meaning in intonation: rising = doubt, question (s. 1.3.4)

meaning in discourse: *How are you? ... Nice talking to you.*

meaning in paralinguistics (intercultural differences!):

- nodding and shaking one's head
- smile + *you know I hate you*

4.4. Lexicology

4.4.1. Introduction

= subfield? of semantics,
investigates the lexicon of a language and its entries
• lexicon entries: not seen as list of isolated elements
lexicology tries to find generalizations and regularities
• lexicon: vocabulary considered from a synchronic, systematic perspective
lexicography = study of dictionaries and dictionary making
• common: assumption: English contains a large central area shared by all speakers (cf. diagram)
• literary: contains scientific, foreign and archaic words
• colloquial: contains dialectal, vulgar, slang and technical elements

4.4.2. Structure of the English vocabulary

English vocabulary:
not homogeneous,
A temporally variable
(changes over time)
• synchronic view vs.
diachronic view
although words look
familiar their meanings
can have changed
B spatially variable

BrE	vs.	AmE
railway		*railroad*
luggage		*baggage*
lorry		*truck*

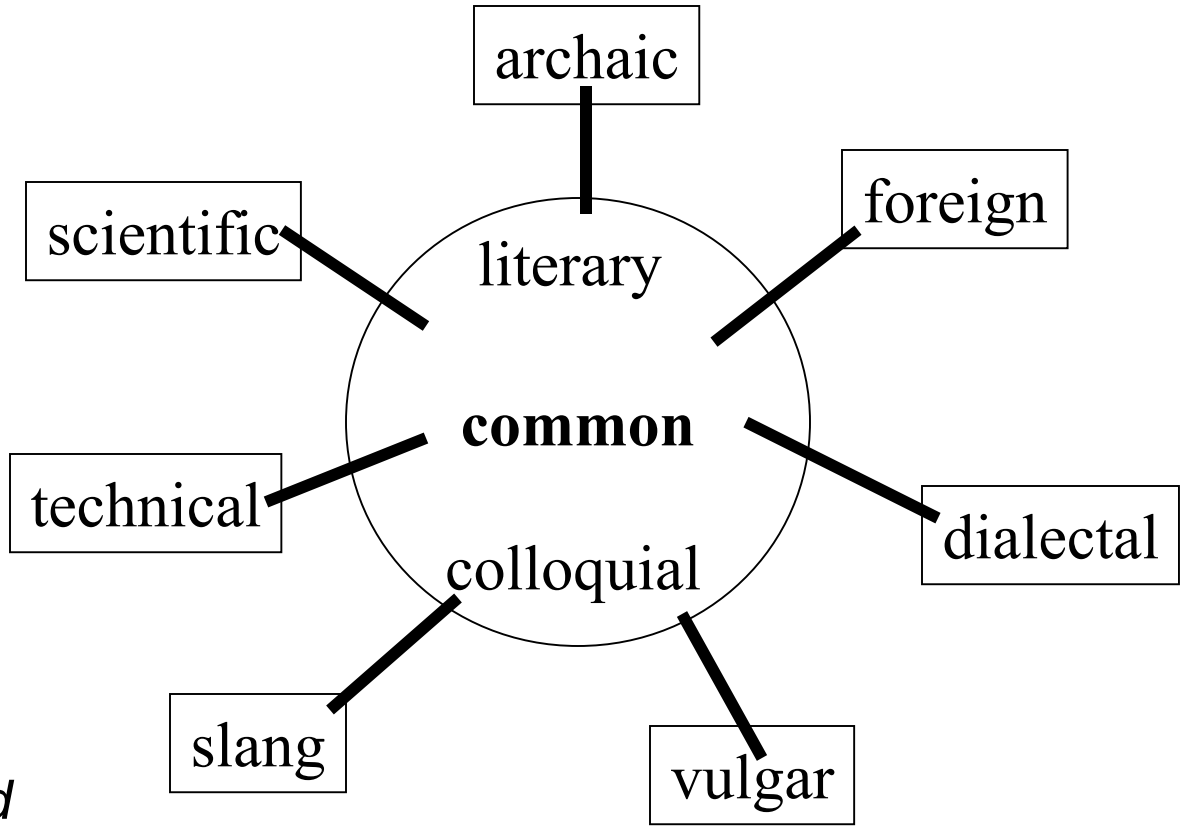


diagram adapted from Lipka 1992

4.4.3. History of the lexicon: etymology

waves of (lexeme) invasions: loan words (and loan translations)

- continental Latin: *cheese*
- Celtic (mainly in place names: -*comb*)
- North Germanic: *skirt* (vs. shirt)
- Norman: *beef*, *pork*, *mutton* (vs. *ox*, *pig*, *sheep*)
- French: *guardian* (vs. *warden*)
- Latin/Greek: *disk* (vs. *dish* Greek via Latin)
- colonial (Dutch): *yacht*
- imperial/global: *pizza*, *curry*, *nasi goreng*

typological parallels German - English
and differences (false friends: *actual*, *brave*, *bright*)

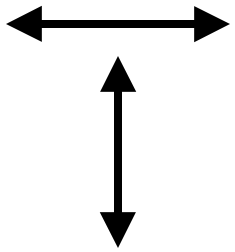
homophones through phonetic change:
meet vs. *meat*, *queen* (vs. *quean*)

types of semantic shifts:
town (*Zaum*), *knight* (*Knecht*), *knave* (*Knabe*)

4.4.4. Fundamental distinctions in lexicology

paradigmatic vs. syntagmatic relationships in a lexicon
language is linear, elements follow sequentially
syntagm: successive linguistic elements that are combined (de Saussure)
paradigm: linguistic elements in opposition or alternative to the same position in a syntagm/sentence
syntagmatic relations = elements can be combined, co-occur
paradigmatic relationships = elements can be exchanged, replaced

He	*can*	*go*	*tomorrow*	syntagmatic
she	may	come	soon	
I	will	start	next	paradig-
You	could	sleep	now	matic



4.4.5. Dictionaries of the English language

bilingual vs. monolingual dictionaries

monolingual dictionaries should contain the following information:
pronunciation, definitions, collocations/ idioms, notes on usage

for English: Oxford English Dictionary OED (**à** COED)

learner dictionaries:

- Advanced Learner's Dictionary of Current English (OALD)
- Longman Dictionary of Contemporary English (LDOCE)
- COBUILD English Language Dictionary
- Cambridge Dictionary of International English (CIDE)

others:

Webster's New World Dictionary of the American Language
on-line dictionaries (e.g. at TUC, TUMunich)

4.5. Pragmatics

4.5.1. Introduction

pragmatics: study of how context influences the interpretation of meaning

context includes: speaker, hearer, third party participants, beliefs, world knowledge

- deals with people's use of language
- is part of performance
- is concerned with principles people use when communicating
- cannot be captured by semantic theory, cf.: *It is rather cold in here*

performative utterances (the saying of the words constitutes the performing of an act) e.g. *I name this ship Queen Elizabeth.*
Speech acts (Austin/Searle): locutionary, illocutionary, perlocutionary acts
e.g. There's a bull in the field.

à cooperative principle (Grice)

4.5.2. Grice's rules of cooperative behavior

1. maxim of quantity
- give the right amount of information when talking
- make your contribution as informative as required but not more

2. maxim of quality
- be truthful, try to make a contribution that is true
- do not say anything for which you lack evidence

3. maxim of relevance
- give a reply that fits the question

4. maxim of manner
- be clear and orderly, avoid obscurity and ambiguity

4.5.3. Implications and facticity

Conversational implicatures: something is understood although it has not been explicitly said -> drawing conclusions from what is said:

He continued writing the essay - implication: *He wrote an essay before*

Facticity of utterance

- factive verbs: situation is true: *The cat is in the garden*
- non-factive verbs: situation has some probability:

I believe the cat is in the garden

- contrafactive verbs: situation is not the case:

I wish the cat was in the garden

- performative verbs: statement is an action itself

I warn you, John accuses Mary, Fred promises ...

-> we act with speech (speech act)

5. Macrolinguistics
5.1. Textlinguistics
5.1.1. Mode: spoken vs. written, a continuum?

	spoken =	**written =**
sentences	short	long
subordination	restricted	complex
sentences	incomplete	complete
constructions	‘ungrammatical’	grammatical
premodification	less	heavy
active form	predominate	also passive
vague language	*thing, stuff*	technical terms
repetition	+ hesitation	+ structured
prefabricated fillers	*you know/see* (hedges)	
deictic (immediacy)	*here, now*	*there, in those days*

difference not always clear-cut (“written to be spoken”)

5.1.2. Coherence and cohesion in texts

(Halliday/Hasan 1985)

- coherence = extralinguistic factors contributing to the creation of texture

- cohesion = linguistic means which create texture

  e.g. anaphoric reference: *this* (*that*)
  substitution of NP by personal pronouns: *he/she*
  conjunctions: *but*, *while*
  adverbs: *first, then, finally*
  lexical repetition:
  incl. synonyms: *problem – difficulty - issue*
  hypernyms: *plant: flower, tree, bush*
  Parliament - Lords

5.1.3. Text-types

text-type classification, e.g. *sermon, cooking recipe*

directive, informative, expressive functions >>

- narrative = novels, reportage (+dynamic verbs, time adverbials)
- descriptive = background reports (+stative verbs, place adverbials)
- instructive = manuals (+imperatives, politeness forms)
- expository = definitions, declarations (+modal verbs)
- argumentative = theses, advertisements (+negations, sequencing)

but few text-types have only 1 function, e.g.

adds are argumentative-persuasive (*this is good because*) + directive (*buy now*)

5.2. Sociolinguistics

5.2.1. Language in multilingual speech communities

5.2.1.1. Language choice

code: code-switching and -mixing (transfer and integration of loanwords)
diglossia: clear functional separation of domains of related (Ferguson) or unrelated (Fishman) language varieties
domains: family, religion, education, law, administration, media, etc.

5.2.1.2. Language maintenance and shift (Gaelic, Welsh)

language death and loss (Cornish), language revival (Hebrew)

5.2.1.3. Linguistic varieties and multilingual nations

vernacular and standard languages
lingua francas: e.g. Latin, Kiswahili, English for Academic Purposes (= EAP)
pidgins and creoles in West Africa (Krio, NP), the Caribbean (Jamaican Creole), the Pacific (Kriol, Tok Pisin)
national (e.g. in Malawi: Chichewa) and official (e.g. English) languages

5.2.1.4. Language planning

(e.g. Kiswahili in Africa, but also in language teaching)
selection + codification + elaboration + securing variety acceptance

5.2.2. Language variation according to users

5.2.2.1. Regional and social dialects

5.2.2.1.1. International varieties

British and American preferences, but also Canada, Australia/New Zealand (*wellies - gummies* [*gumboots*] NZ)
e.g. in lexicon: *luggage - baggage*, *lift - elevator*,
in grammar: *have you got/eaten; do you have/did you eat*?

5.2.2.1.2. Intranational varieties (dialects)

e.g. Geordie speech used in television programmes
(stereotyped notions in creative literature to create character and setting)
famous isoglosses (dialect lines) in pronunciation:
NEngland *but*, *grass*; WEngland/New England: postvocalic <r>
Cockney: systematic vowel shift (*team* > *tame* > *time*),
rhyming slang (*trouble* and *strive* = *wife*)
non-standard grammar less regional, more social (see 5.2.3.)

5.2.2.1.3. Cross-continental dialect chains show the arbitrariness of the terms ‘language’ vs. ‘dialect’

e.g. Scottish: *I won't do it - I'll not do it*

5.2.3. Social variation

Standard English = a dialect with an army and a navy/
with a dictionary and a grammar

a short history of standard English from OE Winchester to the Queen?
vocabulary: U [=upper-class] and Non-U (*lavatory - toilet*)
RP a social accent (through public schools, BBC)
non-standard:
pronunciation: <r> postvocalic (+prestige in AmE, -prestige in BrE!),
<h> aich-dropping/-pronouncing (lower – middle class)
grammar: 0 present tense 3rd person sing,
0 past tense,
double negation (*I ain't seen nothing*)
variation as an indicator of language change in progress (see 2.6.)

famous studies: Labov in NY, Trudgill in Norwich, Milroy in Belfast

5.2.4. Gender

NOT grammatical (=noun classes), referential (*husband*) or collocational (*giggle, handsome guy vs. pretty girl*)

5.2.4.1. Sex-exclusive speech differences (in non-Western communities)

e.g. Japanese *atashi* (=female *I*) - *boku* (=male *I*);
marry (=male) - "get married" (=female)

5.2.4.2. Sex-preferential speech differences (in Western Communities)

e.g. glottalized [p],[t],[k]

women tend to use more standard forms than men

explanations: - (social) status-conscious

- women's role as guardian of society's values
- subordinate groups must be polite (face-protecting)
- vernacular forms express machismo
- others: different responses to interviewers collecting data

(women are more cooperative conversationalists, working-class men react against middle-class academic speech)

5.2.5. Age

(distinguish between language change and age-specific language)

society's expectations:

- younger and older more vernacular forms,
- middle years normal patterns of standard/prestige forms
  solidarity markers

e.g. Bergen Corpus of London Teenage Language (COLT):

500 000 words of 50 hours of recorded conversations from 31 boys and girls aged 13 to 17 from 5 socio-economic groups

5.2.6. Ethnicity

linguistic ideal: "all languages are equal"
sociolinguistic reality: "all languages have resources to be developed/reflect social status of speakers"
other markers: food, dress, religion

5.2.6.1. Black American (the Creole connection)

e.g. absence of copula *be*, multiple negation

5.2.6.2. Black British

an anti-language = opposition to mainstream values

networks:
density (number of contacts) - plexity (multi-dimensional interactions: multi-/monoplex)

5.2.7. Language change

5.2.6.1. Speaker innovation and variation
spontaneous vs. network-specific
by imitation, innovations spread by adoption, diffusion in community networks (weak ties of link person as innovation bridge)
e.g. *which* - *witch*, *whether* - *weather*
Northern Ireland: *man* [mo:n] - *map* [ma:p], *mo'er*
(cf. letters-to-the editor)

5.2.6.2. Social marking postvocalic [r]

5.2.6.3. Spread (wave metaphor)
style to style (casual > formal, incl. spoken > written)
word to word (lexical diffusion): e.g. today: *really*=*rarely*,*fear* = *fair*

5.2.6.4. Reasons
status: particularly upper working class (less consciously) from neighbouring communities with greater social status
sex: men as innovators tend to introduce vernacular, women prestige forms (face-to-face) interactions

5.2.6.5. Research approaches
real-time and apparent-time (=age-grading) studies

5.3. Psycholinguistics

5.3.1. Approaches in psycholinguistics

concerned with psychological processes that make acquisition and use of language possible approaches:

1. language comprehension (spoken and written)
2. speech production
3. language acquisition

language = a *cognitive system* internalized within the human mind/brain (correspondence hypothesis)

• neurological foundations of language: *particular* areas of the neocortex are responsible for human language faculty (results from aphasia research)

aphasia: impairment or loss of language ability due to brain damage

5.3.2. A comparison of L1 and L2 acquisition

children immitate:
differences between pronunciation and grammar, etc.

- innateness hypothesis
- critical age hypothesis

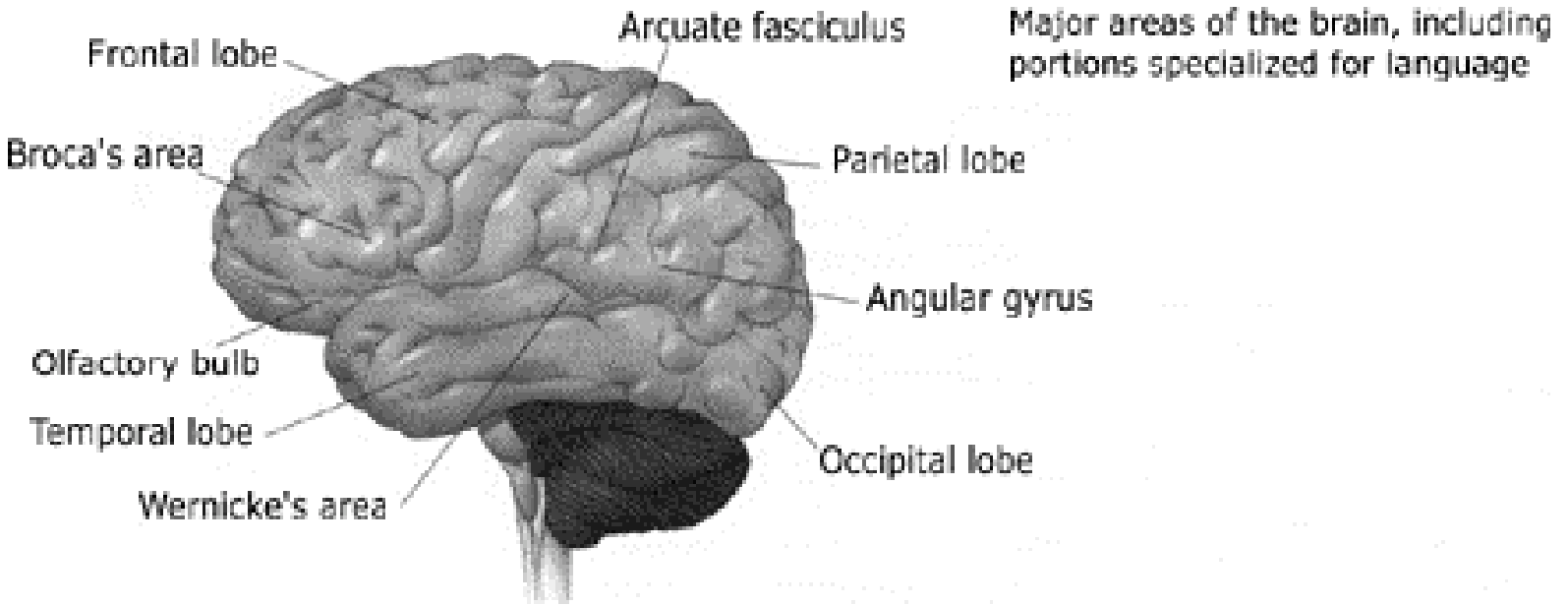
lateralisation = specialization of the brain hemispheres

- monitor hypothesis (Krashen)

interference = negative transfer (="Saxon English")
vs. simplification, overgeneralisation (within L2)
interlanguage = between L1 and L2 (source vs. target language)

5.3.3. Neurological foundations of language

Broca: located lesions in left hemisphere;
related handedness to speech capability
plasticity of the brain (i.e. temporal variability)
Wernicke: separated auditory nerve in the left hemisphere



5.3.4. Language-related areas of the brain

Broca aphasics:	Wernicke aphasics:
• nonfluent • agrammatical • morphemeless • unimpaired comprehension	• fluent (logorrheic) • impaired meanings • neologisms • severely impaired comprehension

- spatial: lateral distribution: - detectable in lesions; PET, fMRI scans
- temporal: brain plasticity; learnability constraints

5.3.5. The paradox of psycholinguistics

L1 acquisition enables children to produce virtually infinite amounts of linguistic data

input (is not intake!) includes:

- distorted input (also: deviant input; Chomsky) can be:
  mispronunciations, slips of the tongue
- omitted rules
  inference of rules out of defective material
- negative evidence
  = pointing at errors

typical errors in L1: **go-ed*
atypical errors: **I no like syntax*

5.4. Corpuslinguistics

5.4.1. Definition

corpus1 = body or collection of written or spoken material upon which linguistic analysis is based

corpus2 = machine-readable
"representative", i.e. stratified "model"
more than a text collection
for computer-based examination

with tools/corpus-analysis software:
WordSmith
Sara (BNC)

5.4.2. Reasons for the popularity of corpus linguistics esp. among non-native speakers

- citations used as a sample of language
- provides a view beyond individual experience
- rules out individual salience
- computer processable

output: - concordances (KWIC=key word in context)
- collocates
- relative vs. absolute frequencies

5.4.3. Corpus search strategies

relative frequency of a word form: standard deviation from mean frequency of word forms

- collocation = the appearance of one particular word form in certain distance of another particular word forms > different meanings can have different collocates

- colligation = the appearance of one particular word form in a particular grammatical structure

- connotation = the semantic environment, can have positive or negative value

5.4.4. Corpus research examples:

- How frequent is a particular morphological form/grammatical structure?
- Which particular structures have particular meanings?
- Which particular structures have particular locations in texts?

corpus tasks have degrees of complexity

- relevance of tagging:
  - parts-of-speech (POS)
  - semantic

5.4.5. Developments in corpus compilation: reference data

1950s American structuralists, e.g. Harris

1959 Quirk: Survey of English Usage (SEU)
1,000,000 words written/spoken 1953-1987
>London-Lund corpus of spoken English

1963/64 Francis/Kucera: Brown Corpus
1M of written American English

1970-1978 Johansson & Leech: LOB parallel to Brown
1M written BritE (Lancaster-Oslo/Bergen Corpus)

1980 - Cobuild Corpus (Birmingham, Sinclair) > Bank of English

1990 - International Corpus of English:
UK, US, EA (KE/TZ), ZA, HK, AZ, NZ, P

1990 - 1993 British National Corpus 100M (10M spoken)

for latest developments see

http://www.tu-chemnitz.de/phil/english/chairs/linguist/independent/kursmaterialien/introling