

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	<i>thing, stuff</i>	technical terms
repetition	+ hesitation	+ structured
prefabricated fillers	<i>you know/see</i> (hedges)	
<i>deictic (immediacy)</i>	<i>here, now</i>	<i>there, in those days</i>

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*

Lords - Parliament

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*),

rhyiming 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/
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 (cf. in New York - Norwich),

<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.)

(Labov in NY, Trudgill in Norwich)

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 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* \rightarrow *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)

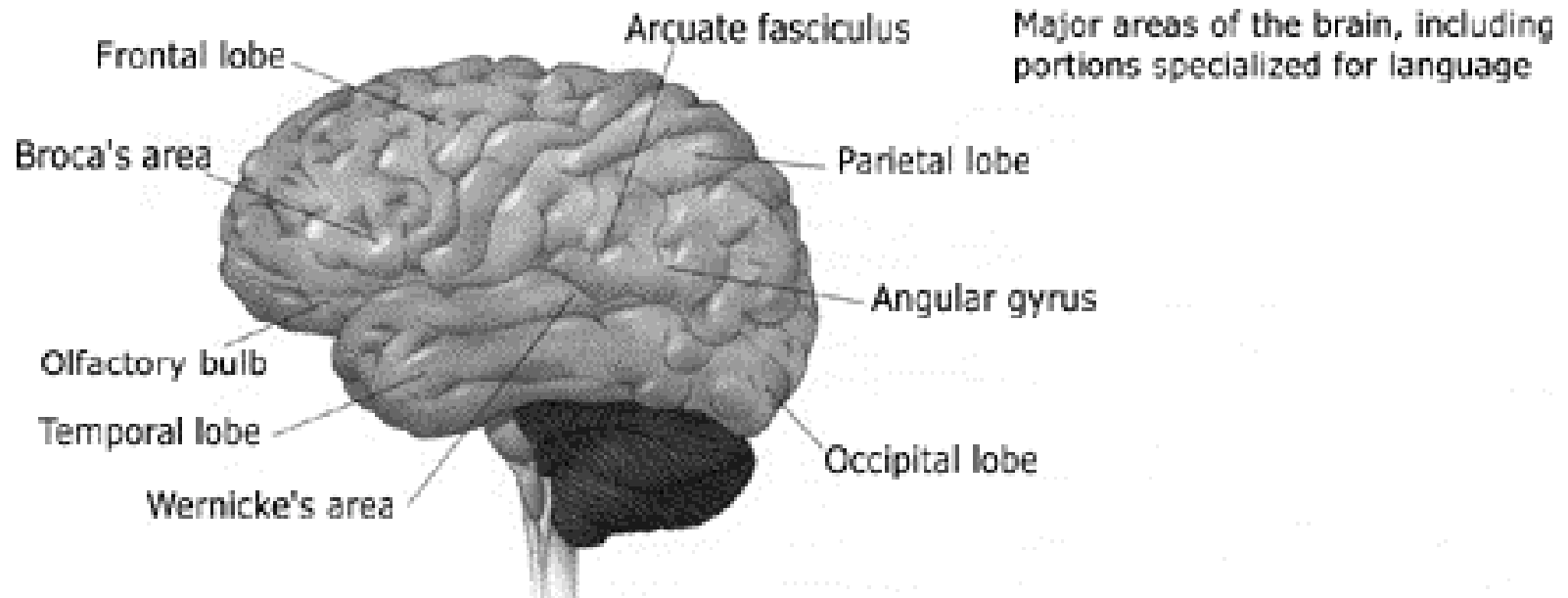
vs. simplification, overgeneralisation (within L2)

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:

- nonfluent
- agrammatical
- morphemeless
- unimpaired comprehension

Wernicke aphasics:

- 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 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 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

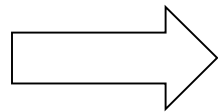
- collocation = the appearance of one particular word form in certain distance of another particular word forms
> different meanings can have different collocates
- colligation = the appearance of one particular word form in a particular grammatical structure
- connotation = the semantic environment, 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 - 199x 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/introoling>

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	<i>thing, stuff</i>	technical terms
repetition	+ hesitation	+ structured
prefabricated fillers	<i>you know/see</i> (hedges)	
<i>deictic (immediacy)</i>	<i>here, now</i>	<i>there, in those days</i>

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*

Lords - Parliament

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*),

rhyiming 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/
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 (cf. in New York - Norwich),
<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.)

(Labov in NY, Trudgill in Norwich)

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 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* \rightarrow *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)

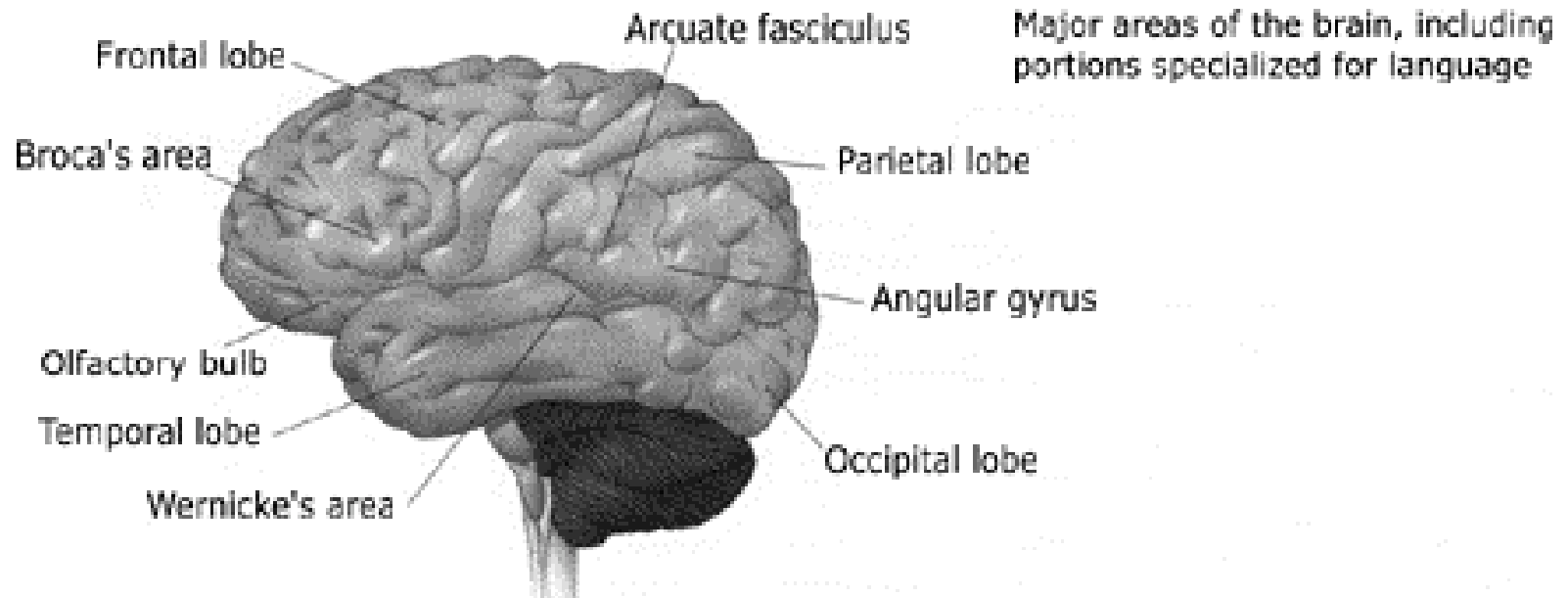
vs. simplification, overgeneralisation (within L2)

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:

- nonfluent
- agrammatical
- morphemeless
- unimpaired comprehension

Wernicke aphasics:

- 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 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 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 - 199x International Corpus of English:
UK, US, EA (ke/tz), ZA, HK, AZ, NZ, P

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