Glossary of Linguistics

Conventions in this document.
Grammar rules look like:
Base[P][#][Suffix]

Grammar rules start with the part of speech, followed by

1. An optional P, indicating a (possibly infinite) phrase
2. An optional place number (1..4)
3. Followed by optional 2 letter suffixes indicating the kind of noun phrase:
   - SC: can go with a single, countable noun phrase
   - RL: can go with a plural, countable noun phrase
   - UC: can go with a uncountable noun phrase
   - CN: can with a common-noun phrase
4. Phrase rule names have suffixes to indicate the type of function they build:
   - _itr. They take no arguments and are iterators.
   - _w: They take no arguments and a return value
   - _LZ. They take an argument in and return a value
   - opt_: Used when the phrase is optional (similar to the ? quantifier in regular expressions.) Match an explicit phrase once or a NIL phrase.

In addition to the phrase rule, grammar rules include a *denotational signature* and *intensional rule.*

Data structures

- trie
- table
- phrase grammars

absolute words absolute words are not subject to comparison

accommodation Shifting (esp unconsciously) ones speech patterns, dialect, enunciation to adopt the patterns from others.

adjectives Descriptive. Limiting.

There are four major kinds (zones) of adjectives, each placed relative to each other

Note determiners: The determiners have taken on a role traditional grammar reserved for adjectives.

see also *article*

absolute Words (e.g. supreme, infinite) that there can't be "more" of.

comparative May have a generic form, comparative form, and a superlative form. Regular adjectives have a comparative form by adding –er, and a superlative form by adding –est.

The intensional formula for a comparative is found by

1. Determine if the adjective is definite or indefinite
2. Look up the adjective forms (definite from table 3, indefinite from table 4)

3. Determine if the adjective is negative or positive
4. Look up the formula for the positive (or negative) form
5. The comparative form is the difference form applied to
   $\lambda_f\ \lambda_{right}\ \lambda_{left}$[f(right)(left) > 0]
6. For indefinite adjectives:
   a. Apply the TBD form to formulae for the adjectives generic definition
   b. Apply comparative forms to the formulae for the adjectives comparative definition
   c. Apply the additive to the formulae for more comparative definitions
7. For definite adjectives
   a. Apply to the generic form
   b. Apply the equative to the formulae for more generic definitions
   c. Apply comparative forms to the formulae for the adjectives comparative definition
8. Finding a ranking function for the adjective.
   a. Definite adjective's might be looking up a property (e.g. mass) for an item; such a function can be found by applying the formula:
      $\lambda_{propertyName}\ \lambda_{item}$[property(item, propertyName)]
   b. Indefinite adjectives do not have a ranking function
9. Look up the formula for the positive (or negative) form and apply the ranking function

The intensional formula for the superlative formula is found by applying the comparative formula to:

$\lambda_{cmp}\ \lambda_{right}\ \lambda_{left}$[Ø={x s.t. x ∈right ^ (x≠left) ^ ¬cmp(left,right)}]

Table : *Some example irregular forms*

generic	*comparative*	*superlative*	
good / well	better	best	indefinite
little	less / lesser	least / littlest	indefinite
much	more	most	indefinite

Table : *adjective forms*

form	*Difference*	*Equative*	*Additive*
negative	$\lambda_{left}\ \lambda_{right}$[right - left]	$\lambda_{left}\ \lambda_{right}$[left ≅ right]	$\lambda_{left}\ \lambda_{right}$[left - right]
positive	$\lambda_{left}\ \lambda_{right}$[left - right]	$\lambda_{left}\ \lambda_{right}$[left ≅ right]	$\lambda_{left}\ \lambda_{right}$[left + right]

Table :*definite adjective formulae*

$\lambda_f\ \lambda_p\ \lambda_{right}\ \lambda_{item}$	[right (f(p(item))) (p)]	taller than bill

Table : *indefinite adjective formulae*

$\lambda_f\ \lambda_{right}\ \lambda_{left}$	[f(left) (right)]	more $20
$\lambda_f\ \lambda_{right1}\ \lambda_{right2}\ \lambda_{item}$	[right2 (f(right1(item))) (right1)]	more money than bob
$\lambda_f\ \lambda_{right1}\ \lambda_{right2}\ \lambda_p$	[right2 (f(p(right1))) (p)]	more cats than dogs
$\lambda_f\ \lambda_{right1}\ \lambda_{right2}\ \lambda_{item}$	[right2 (f(right1(item)) (item)]	more money than sense
$\lambda_f\ \lambda_{right}\ \lambda_{item}\ \lambda_p$	[right (f(p(item))) (p)]	more than bill

demonstrative	see *pronoun (demonstrative)*
indefinite	all, some, any, none
interrogative relative	whose, which, what
numeral	two, first, etc.

ordering	The royal ordering adjectives is:
	▪ Determiner
	▪ Observation (e.g. beautiful, delicious)
	▪ Physical description, in order:
	  ○ Size
	  ○ Shape
	  ○ Age
	  ○ Colour
	▪ Origin
	▪ Material
	▪ Qualifier (e.g. touring bus, hat box, etc.)

possessive	my, his, her, your, our, their

Figure : *Adjective rules*

```
adj1 ::= absolute
adj1 ::= intensifying e.g. newest
adj2 ::= other
adj3 ::= participles
adj3 ::= color adj
adj4s::= adj4
adj4s::= adj4 conj adj4s
adj4 ::= adj4F
adj4 ::= noun4G
adj4F::= sound
adj4F ::= noun4G
```

Table : *article words*

binding	*words*
	both
adj_SC	every each either neither
adjA_UCPC	all (indefinite)
adjB_UCPC	some any must a-lot lots enough (indefinite)
adj_SCUC	this that (demonstrative)
adj1_UC	least little much less

adverb

modify things other than nouns – adjectives, verbs, adverbs, clauses, sentences. By scanning for adverbs one can answer some of the how/who/what/when/extent/what way/etc questions. They have the form

$$\lambda_{right1}\ \lambda_{subject}\ \lambda_{object}$$

The subject (right) may be anaphoric.

Adverbs of degree apply a transform to the object. These adverbs consume adverbs to the right.

Table : *adverbs*

word	*comparative*

as	$\lambda_{right1}\ \lambda_{right2}\ \lambda_{left}$[right1(left) = right2(right1)]
first	$\lambda_{n}\ \lambda_{clause}$ [x s.t. clause[0…n-1]]

Table : *adverbs of degree*

word	*comparative*	*c*
almost	$\lambda_{right}\ \lambda_{object}\ \lambda_{subject}$[right1(c·object)(subject)]	c < 1
atleast	$\lambda_{right}\ \lambda_{object}\ \lambda_{subject}$[right1(c·object)(subject)]	c = 1
much	$\lambda_{right}\ \lambda_{subject}$[right1(subject) > c]	c > 1
nearly	$\lambda_{right}\ \lambda_{object}\ \lambda_{subject}$[right1(c·object)(subject)]	c < 1

ordering

The 'royal' order of adverbs is:

- Verb
- Manner
- Place
- Frequency
- Time
- Purpose

adverbial

A modifier – see the kinds of modifiers

position

Positions adverbial can appear in the declarative form of a clause:

- Initial position (i.e., before the subject)
- Medial position #1: (a) immediately before the operator (DO, etc.), or (b) between two auxiliaries.
- Medial position #2: (a) immediately before the verb, or (b) before the complement in intensive BE-clauses, e. g. "He is soon to be transferred."
- End position: after an intransitive verb, object or complement.

intensifier

Adverbial Intensifier is an adverb that qualifies the intensity of the Quantifying Pronoun, Adjective, Adverb or Verb adjacent to it.

May be subclass of a degree adverb; no consensus on this.

agreement

Changing the form of word based on the other words it relates to. This is often grammatical properties and orthography to pair with main element.

ambiguity

Uncertainty about which of a word's or expression's possible meanings is the one intended. *Note: ambiguity may not be noticed, and the intended meaning may be missed as a result.*

See also *diaphoric, modifiers (vague, dangling), vagueness*

grammatical ambiguity

The grammatical structure of a sentence allows it to be understood in more than one way; it is not clear from the context which understanding is the intended one. Often surface structure ambiguity.

lexical ambiguity

A word or expression has more than one meaning sense, it is not clear which to employ.

referential ambiguity

An expression of reference can be interpreted as designating more than one thing. Can be semantic structure (e.g. "every man loves exactly one woman").

fallacy of equivocation

Treating two distinct meanings of a word as though they were the same.

example

"We painted the walls with cracks"

- The wall had broken (cracks) and it was painted over
- We painted on representations of cracks

anaphora See *reference*

archetypical language understanding assumptions

Michael Kac

1. The source is in text form rather than spoken or other form
2. Understanding is broken down into distinct analysis types – i.e. syntactic, semantic, usually in that order
3. The syntax is processed in a single unidirectional pass, usually left to right
4. The syntax process is formed modularly (e.g. interpreter, compiler) in a fashion independent of the source language (i.e. the grammar is supplied separately)
5. The grammar is formed as a context-free phrase structure grammar, possibly with extensions.

argument graph

An argument has an introduction, thesis, grounds, example, refutation of opposing arguments.

Types of belief contents: fact (corresponding to the real world), rule (causal relationship between two states of affairs), judgment (Good or bad).

Nodes have labels like: thesis, conclusion, head, reason, anti-opposing-constituent, cause of thesis, example of thesis, deny.

Node types: issues, positions, arguments.

link types: responds-to, questions, supports, objects-to, specializes, generalizes, refers-to, replaces.

argument structure

Jane Grishaw, Argument Structure

represent the relevant syntactic inputs for interpreting a verb.

part of *linking*

articles

Articles go before a noun

indefinite articles

Designates one of a group but not specific.

Table : *article combining rules*

binding	
art_SC	a an
art_CN	the
SCUC	this that

Table : *Articles*

	number	
indefinite	singular	a / an
definite		the

attachment
minimal

Minimal attachment: [s [np The man] [vp [v kept [np the dog] [pp in the house]]]

Right association: [s[s [np The horse] [vp [v race] [pp [p past] [np the barn]]]] [vp [v fell]]]

automation

https://www.projectoxford.ai/
https://www.gallerty.cortana analytics.com
http://www.alchemyapi.com

spell check

attitude

transcription: speech to text

text to speech

speaker recognition

topic classification

language understanding

extract: entity, sentiment, analysis, concepts, which language

blackboard parsing

At the main clause level, identify main clause's finite verb and other single-word elements

Identify complex sentence elements – subject, object/predicative, adverbial, done heuristically.

capitalization

Rules of

case

a grammatical property, re the role of the noun.

see also *mood.*

Table : *Example of noun cases*

Case	*Role*
ablative	Used after preposition meaning from, indicates direction or agent hood
accusative	Noun is object in the clauses; must be used after preposition meaning 'to'
dative	Noun functions as indirect object in clause, indicating benefit
genitive	Item referred to by name, is possessor of something
nominative	noun is subject in the clause
objective	accusative and/or dative cases
possessive	indicates relationship of possession
reflexive	
subjective	another name for the nominative case
vocative	Form used to address someone, exclamations

cataphora

See *reference*

classes

Substance & physical objects, quantity, quality, relation, place, time, position, state, action, affection, events, ideas, concepts, plans.

common usage

Field linguistics idiom. When a term becomes frequently used for a broader activity than the than the trademark holder would prefer the risk losing their trademark.

Demonstrate that Xerox is never used to refer to copies or the process of copying; Kleenex is never used to refer to nasal tissue; Coke is never used to soda pop; and "phone" is solely used to refer to the specific Alexander Bell devices sold (and never the process of using analogous items or a phone network).

communication classes of topics

Motion, possession, perception, communications, competition, change, cognition, consumption, contact, creation, emotion, bodily care & function, social behaviour and interactions. Expressions of weather.

http://www.cogsci.princeton.edu/

See also *dialogue*

visual communication method choice of information type and communicative function

information types

concrete, abstract, spatial, covariant, temporal, quantification, negation

communicative function

attract-attention, compare, elaborate, enable, elucidate, label, motivate, evidence, background, summarize

concessive clause

see *adverbial, modifier*

conclusion

as a result, as is implied by, accordingly, consequently, hence, I conclude that, In

indicators consequence, it follows that, seeing that, so, proves that, therefore, thus, we may infer, which allows us to infer, which entails that, which implies that, which means that, which points to the conclusion that, which shows that

conjunction conjunction ::= X "and" X

One for X = NP, VP, S, PP, ADJP, ADVP, N, V, Aux, Adj

types Conjunctions of addition and replacement

Conjunctions of comparison, contrast and concession

Conjunctions of exemplification and restatement

Conjunctions of cause and condition

Conjunctions of time

contractions rules and limits of. [] [] -> no. Idiomatic.

see also *flexion*

copula A word that links the subject with a predicate. Usually is a verb. The predicate may be:

- noun or noun phrase
- adjective or adjective phrase
- a prepositional phrase
- adverb or adverbial phrase expressing time or location

The statement may express

- Identity: The subject and predicate have the same referent, or express an identical concept
- Membership of a class, esp. that the subject is a member of the class referred to by the predicate
- Properties (e.g. that the subject has the property referred to by the predicate)
- Relationship
- Position

date

```
Date_preposition_phrase:= Date_noun_phrase
Date_preposition_phrase:= Date_noun_phrase "after" Date_preposition_phrase
Date_preposition_phrase:= Date_noun_phrase "before" Date_preposition_phrase
Date_adj::= "this" | "last" | "next" | "this" "coming" | "the" "previous"
Date_noun_phrase:= Date_adj Date_noun
Date_noun_phrase:= "today" | "tomorrow" | "now"
Date_noun:= "week" | "month" | "year" | "decade" | "century" | "quarter" |
Day_of_week|month | "fortnight"
```

definition A definition is an explanation of a words meaning and use.

reportive definition Definitions intended to explain how the words are actually used (in the 'field').

disciplinary A report about the way a word is used in a particular discipline or specialty area

historical A report about how a word was used during a particular historical period. May be lexical or disciplinary.

lexical a report about the way a word is used in every day life

précising Restricts the ordinary meaning of a word to make the meaning more exact in a certain context.

stipulative A statement of the rule that will be followed in used the defined word. A resolution to use a word in a certain way, to assign the word a particular meaning. (Cannot be true or false).

synonym Providing words with similar meaning

genus and species Mentioning a feature of an object a word refers to that places the object within a

class, then mentioning another feature that places the object within a subclass.

enumeration — Listing all the items to which the word refers. Most words belong to an unlimited class, making this approach unacceptable in the general case.

ostensive — giving examples of what the word may refer to.

term definition — Concerned with how to represent the meaning of a term. Langacker worked out a system of dividing the definition into a base (general) and profile (specific to the term). This is similar to prototype object system, but I do not know if he includes a system to refer to the context network of inheritance. Also interest with polysemy.

deixis — Interpretation requires context information to know which perspective to interpret it from. Types: spatial, temporal, social, discourse

determiners — can be definite, or indefinite. Many forms.

The concept of determiner phrase is not used in traditional grammar; it does much of that adjectives did in those grammars.

some I don't know about yet: another (a whole nother, anoda), any old, atta (that's the, that's a), beaucoup, certain, dat, dis, fuck all, last, next, nil, overmuch, own, quodque, said, umpteen, various, wat, yon, yonder

Table : *determiners*

		terms
additive	?	more (more and more)
cardinal numbers	definite	zero, one, two, fifty, infinite, ...
alternative	either	another, other, somebody else, different
articles	either	a, an, the *see article*
degree /partitive	either	many, much, few, little, couple, several, least, most (mos'), not a little, a number of, quite a few, fewscore
demonstratives	definite	this, that, these, those, which *see also pronouns (demonstrative)*
disjunctive	indefinite	either, neither
distributive	either	each, every, each and every
elective	indefinite	any, either, whichever
equative	definite	the same
evaluative	definite	such, that, so
exclamative	definite	what *eyes*!
existential	indefinite	some, any (anny, eny)
interrogative and relative	indefinite	which, what, whichever, whatever
maximal	?	the most
minimal	?	the least, fewest
multal	?	a lot of, many (many a), several, much
negative	indefinite	no, neither
paucal	?	a few, a little, some
personal	definite	we *teachers*, you *guys*
possessive	definite	my, your, our, his, her, *etc*. *see possessive pronoun*
qualitative	?	that, so
quantifiers	either	all, few, many, several (severall), some, every (hevery, euerie, everie), each, any, no (nary a), not even one, none, etc.
subtractive	?	less, fewer
sufficiency	?	enough (enuf, enuff), sufficient, plenty
uniquitive	definite	the only, the, this, that, these, those

universal	indefinite	all, any and all, both (bothe)

Table : *quantifier determiner*

term	*formulae*
all	$\lambda_{right}\ \lambda_{p}$ [right = p(right)]
every	$\lambda_{right}\ \lambda_{p}$ [right = p(right)]
few	$\lambda_{right}\ \lambda_{p}$ [|right|·c_0 < |p(right)|]
many	$\lambda_{right}\ \lambda_{p}$ [|right|·c_0 = |p(right)|]
most	$\lambda_{right}\ \lambda_{p}$ [|right|·c_0 = |p(right)|]
no	$\lambda_{right}\ \lambda_{p}$ [Ø = p(right)]
several	$\lambda_{right}\ \lambda_{p}$ [|right|·c_0 = |p(right)|]
some	$\lambda_{right}\ \lambda_{p}$ [Ø ≠ p(right)]

determiner phrases

Determiners tell us which type of noun to expect, and how to approach it. Composed of an optional pre-determiner, a determiner, and an optional post-determiner

see also *pronoun for possessive and demonstrative determiners for english*

number

$\lambda_{right}\ \lambda_{p}\ \lambda_{subject}$ [⟦*num*⟧ ≅ right(p)(subject)]

$\lambda_{right}\ \lambda_{subject}$ [⟦*num*⟧ ≅ right(subject)]

demonstrative determiner

When a demonstrative pronoun is used before a noun.

f:(individual)*

{x | x ∈ *noun phrase* ∧⟦*demonstrative*⟧(x)}

Figure : *Determiner phrase rules*

DetP_SC NounP_SC
DetP_SCUC NounP_SC
DetP_SCUC NounP_UC
DetP_UC NounP_UC
DetP_PLUC NounP_UC
DetP_PLUC NounP_PL
DetP_PL NounP_PL
DetP_CN common noun

det_SC ::= no_predeterminer adj_SC not_poss?
det_CN::= predeterminer art_CN postdeterminer
det_SC ::= art_SC postdeterminer
det_SCUC ::= predeterminer adj_SCUC postdeterminer
det_UC ::= adj1_UC
det_UCPC ::= predeterminer_UCPC determiner postdeterminer
det_UCPC ::= adjB_UCPC postdeterminer
det_PC ::= predeterminer determiner adj_PC

Table : *Examples of each type of determiner*

binding	
Det_SC	a an each either neither every adj1_SC
Det_SCUC	this that
Det_UC	least less (a) little, much
Det_UCPC	all any enough a lot lots most some
Det_PC	a few fewer fewest both many several those these <number>
Det_CN	this, no, possessives, wh-words

central
central ::= art
central := demonstratives
central := possessives
central ::= quantifiers

predeterminer
preDet::= quantifiers
preDet := all both double half twice actions
predeterminer_UCPC ::= adj_UCPC

post determiner
numbers quantifiers (many several)

dialect
patterns, structures, usages, processes, histories

dialogue
See also *communication (topics)*

Dialogue techniques

- Conjunctions
- Change topics
- Details
- Direct statements
- Discuss personal life
- Discuss relationship
- Feelings
- Intensifiers
- Interjections
- Interruptions
- Maintenance statements
- Problem solving
- Questions to stimulate conversation
- Relating experiences
- Task statements
- Tentative statements

Table : *Distinctions in male and female approaches to dialogues*

	Male	*Female*
Discuss personal life in business	Less likely	To establish more business relationships
Disclose personal information	Less likely	More likely
Relationship formation		Discuss personal life
about relationship	Less likely to discuss, and focuses on: what they did, where they when	More likely to discuss, focusing on feelings
express intimate feelings	Less likely	More likely
statements	Direct, declarative (it is a nice day)	Indirect, tentative statements with tag endings, or upward inflection statements (e.g. questions)
intensifiers	Less likely	More likely: few, so really, much, quite
question to stimulate conversation	Less likely	More likely
changing topics	More interjections	More conjunctions
interrupt	Interrupt others, allow few interruptions	Allows more interruptions
rejection	Insensitive	sensitive
problems/troubles	Try to solve	Try to match by relating similar experiences
help	Less likely to ask, accept	More likely to ask, accept
awareness detail	Less likely to ask, accept	More likely to ask, accept
orientation	Task (what are going to do)	Maintenance (are we ok?)

discourse

Discourse model:

- The event a reference refers to.
- The relative temporal ordering of the events.
- Descriptions needed to distinguish the referent.
- Descriptions that relate it to other parts of the conversation.

Two levels of focus of attention.

- Global stack of focus spaces (affects definite description); each stack holds representations of entities.
- Local, centers and centering (affects interpretation of pronouns). The current focus is a backward –looking center. There are many forward-looking centers (potential new focuses)

discourse repairs

Three types of speech repairs: *Heeman and Allan*

1. Fresh starts – the current utterance is abandoned and started over fresh
2. Modification repairs – modifies earlier statements or fragments or whatever
3. Abridged repairs – the repair consists solely of a fragment and/or editing term

"each"

can be used as a subject; takes a singular verb or pronoun

English

English has, at most, three cases: accusative, dative, nominative. Use the French case rules rather than the prescriptions based on Latin. This is recommended since French is a case language that is much closer to English than Latin. This is controversial since most writing guides such rules that are based on Latin.

commands
Verb of action
Direct object
"!" or "."

declarative
Subject
Verb of being
PredictiveNominative
" "

question
Verb of being.
Subject.
Predicate nominative
?

subject
Article Adjective Noun PrepPhrase

prep phrase
Preposition Article Adjective Noun PrepPhrase

verb of action
Adverb verb adverb

clause::= subject VP CR2
CR2 ::= object OAC | OAC
subject ::= you |
OAC ::= object | A | complement
CNP1 ::= RelativePronoun VP
CNP1 ::= DetP? PreMod? Head PostMod?
PostMod = Participle | Prepositions

event

An event changes something, and may specify *van Lambalgen, Michiel; Fritz Hamm* The Proper Treatment of Events

- The initial state of the world
- Who/what initiated the event
- Who/what terminated the event
- What happens during the event

An event may change, with different levels of granularity:

- "the sphere of behavior between verbal, social, and intellectual

- "the predominant part of the body
- "the physical direction of behavior
- "the object of the behavior
- "the behavior setting
- "the tempo of the activity"

"Events may have to be treated different, depending on whether they are considered to be 'foreground' or 'background'"

"in discourse, the order of events is constructed from, e.g, causal information, and not just the order of the sentences; in a sense the temporal relations are derived from such information."

exophora See *reference*

flexion Changing of a word according to its situation:

- Inflection: variable word ending
- Agreement: inflection is change to match another word. Gender, number, case, etc.
- Affix: part of a word which is attached at beginning, end or middle. Indicates gender, number, etc.
- Enclitic. Affix meaning "too" or "and"
- Agglutinization: composing many affixes into a single word.

see also *morpheme contraction*

structures use of a trie to help build this

flexional language "Any language which expresses grammatical relations of words and shades or modifications of their meanings by affixing prefixes or suffixes to the roots of the words. This term thus applies to both aglutinal and amalgamating languages." *Pei*

Fog index analyzing text to estimate how difficult it is to read.

functions of language

- Referential, descriptive statements, deictic
- Expressive (emotive), does not alter denonatives meaning, interjections
- Conative, engage the addressee directly (vocatives, imperatives)
- Poetic,
- Phatic, the parts of language for the sake of interaction
- Metalinguistic, the parts that discuss language, its use

garden path examples

The cotton clothing is made from comes from Mississippi.
The horse raced past the barn fell.
The raft floated down the river sank.
The man who whistles tunes the pianos.
Have the students who failed the exam take the supplementary.
The man who hunts ducks out on weekends.
The prime number few.
Fat people eat accumulates.
The tycoon sold the offshore oil tracts a lot of many want to kill JR.
The old man the boat.

Human beings can and do parse in a deterministic fashion. The garden path effect is evidence of this. Used simultaneous access to 3 words, to avoid garden-pathing where humans avoid it.
Marcus grammar is written in an English-like language called pidgen. This is compiled into LISP. *Mitchel Marcus,* A theory of syntactic recognition for Natural Language *1981, MIT Press.*

Active Node stack;
Constituent buffer (3 items) – either a word or a fully parsed constituent.
"John lifted a hundred pound bags"
"Have the kids running laps been doing it long?"
"Have the kids running laps by the time I get back."

gerund -ing on verb to create a noun form

glossary "collection of terms limited to special area of knowledge" (v1n1). USSID 412 directs forming a glossary per area; 1954: NSA traffic analysis glossary Interim Report #168-54, 1955/1964 Radio Traffic Analysis Manual; 1971 June Basic Cryptologic Glossary.

see *definition*

terminological information (v1n1) semantic unit, source(s), definition(s), significant context(s), field(s) of application, author of information unit

glottochronology

government That which controls the binding of a symbol. Exists over a domain.

see also *binding*

grammar Syntax grammars were developed to predict whether native speakers would accept a sentence. Most *formal* grammars of syntax employ some type of short-term memory, and connection to semantics. Most employ a *simplified and minute* model of semantics: number, thematic role, distinction between thing and action, etc.

case The *who-what-when-where-how-with-why* role a noun or pronoun plays, e.g. actor, instrument, force, recipient. In case-based languages, the pronoun's conjugation signals the case. There is no agreed upon universal set of cases, number of cases, distinction between cases, or refinement of a broader case. Strict syntax, in other languages, provides the same role.

grammar types

- Transformational
- Transition Networks
- Semantic models
- Dependency Theory
- Q&W system
- Formational Theory
- Predictive Analysis
- Statistical Analysis
- Wayne State University MT program

interpretation Many of the formal parsers are too strict, and encounter problems with native text. Practical parsers might drop words that confuse it (in an attempt to recover). Another parser technique when the phrase-structure parse fails:

1. Scan string of words for the subject
2. Then scan for verb
3. Then for object
4. … etc.

A few techniques employ a 'theorem proving' like modification to the phrase structured grammars. Simplified, it works by testing that 'there is a word after this point with the part of speech *XYZ'*.

grammatical induction Given a sentences labeled with part of speech, approximate the grammar

grammatical properties Divided between nouns and verbs

animate (people and higher animals), inanimate (plants, lower animals, machines, things)

aspect: flow of time, or independence (same as tense in Germanic languages); ongoing vs done

Gender: masculine, feminine, neuter

Mood: divided by noun and verb

Number: definite singular, indefinite singular, plural ;

person: first, second, third

polarity: affirmative, negative

tense: past, present

voice: active, passive

Pronunciation

see also *flexion, mood, noun (grammatical property)*

encoding — To make it reasonable to encode this stuff and match, each is assigned 2 bits in a word. Then a mask is employed to track which properties are relevant/ bound.

Heap's law — $Vocabulary\ Size = Kn^{\beta}$

$\beta = 0..1 \quad (0.4 - 0.6)$

$K = 10..100$

$n = \#\ words\ in\ text$

This can be derived from Zipf's law

see also *lexical memory, Zipf's law*

hyphenation — inserting a dash. Seldom needed. Rules about how to do so for each language.

imperative — type of mood which many have rules of conjugation, grammatical number and person., see also *mood*

indefinite number — Function as a number grammatically, but lack a definite quantity. They may be able to compare the quantity. Examples: lots, many, several, some. Note idiomatically some specific numbers are used as indefinite: 101, 40 days & nights, 1001 uses for…

indefinite hyperbolic number — Numbers used to refer to ridiculously large quantities.

affixes — suffices, such as –illion are used to indicate that they are large

infinitive — verb form

inflection — different in form based on context, including spelling of words around it. Varaitions may include flatten, dialect, enunciate

Kaplan context — denotation/interpretation with respect to

- context & world
- Who is speaking
- When
- Where
- In what world
- Assignment of variable (value g)

context set — A set of world

A conversation

Common-grounds: shared information in a conversation

lexical analysis — Ability to assess linguistic complexity is a concern. The methods of lexical analysis include:

- Statistical – pattern classifiers
- Probabilistic
- Geometric
- Discriminant based

- Clustering

See also *grammar, Heap's law*

Levenshtein distance

$$D(a_i,b_j) = \min \begin{cases} D(a_{i-1},b_j) + w(a[i],0) \\ D(a_{i-1},b_{j-1}) + w(a[i],b[j]) \\ D(a_i,b_{j-1}) + w(0,b[j]) \end{cases}$$

Where

- $D(a_i,b_j)$ The string distance from string a of length I to string b of length j
- w(a[I],0) The cost of deleting character a[I]
- w(a[I],b[j]) The cost of substituting a[I] with b[j]
- w(0,b[j]) The cost of inserting character b[j]

library science organize, classify catalog index, archive, retrieve – esp these for media, such as books, movies, art, etc.

resource management.

linguistic data University of Pennsylvania, Penn Tree Bank.

Often count the incidence of a word for a given part of speech.

Brown corpus Brown University. Aged corpus. Kucera.

linguistic drift Semantic shift: semantic broadening or semantic narrowing

Grimm's Law; sound shifts, consonant shifts.

Loss of case affixes in English.

Forms lost for case, number, combinations

linguistic model finite, well-defined vocabulary, composed a large # of coherent sentences

linking how a verbs arguments are linked to syntactic positions in a tree

modal auxiliary verbs I can not eat:

- I have the ability to go without eating
- I am not able or allowed to eat

modality "The attitude of the speaker to the action indicated by a verb" (wikipedia)

commissive commitment to do something, incl promises and threats

deontic how things out to be. divided into commissive modality, directive modality, and volative modality.

directive command, request

dynamic indicates subject's internal capability or willingness

epistemic The "speaker's evaluation of, degree of confidence in, or belief of the knowledge upon which [it] is based."

volative wish, desire

modifier optional. see adverb/adverbial clause, adjective/ adjectival phrase, quantifier; intensifier, description.

attributive: part of the phrase that they modify

predicative: linked (eg a copula) to the phrase that they modify

Can indicate some degree of quantity (relative, or absolute), or it can indicate a degree of truth, usually about a quality.

- Ranks or orders; this may be of items property values, against a ranking, or of a certainty of true of the statements
- spatial or place:

see also *adjective, adverbial*

ambiguity — unclear which element is to be modified

dangling — the element being modified is not in the sentence

postmodifier — A modifier placed after the head

premodifier — A modifier placed before the head

Table : *Types of modifiers*

type of clause	*modified by*
adjectives	adverb
adverbs	adverb
clauses	adverb
determiner	quantifier
noun	adjectives, adjectival clauses, adjectival phrases, article, determiner; noun adjunct
sentence	adverb
statement	adverb
verb	adverb

Table : *Kinds of modifier clauses*

Class	*clause*	*Role*
absolute, superlative	maximizer	Words (e.g. supreme, infinite) that there can't be "more" of.
	booster	Very intense, but there is the possibility of it getting even more intense.
comparative		comparison of some subjective rank or objective value
truth	Approximators	Showing "almost but not quite." Indicating that a statement is near to correct, but not 100% so.
	Compromisers	Allowing opportunity for someone else to disagree with the statement.
	Diminishers	Showing the statement is true to a small degree.
	minimizer	Indicates the statement is not true or true to a very small degree
condition		Possible or counterfactual situation and its consequences
concession		Contrasts multiple statements, suggests the opposite of the main part; may provide surprise.
manner		Discusses someone's behaviour or the way something is done
place / spatial		indicates: location or position. may indicate places (or same place as an event or object); the space may be analytic or topological. In the case of analytic space, entities are mapped to their place in that world at that time and applied to the operation with the space algebra. In the case of topological

	space, the objects are applied to the relation that is looked in that world at that time.
purpose	Indicates: purpose of action
reason	Indicates: the reason for something
results	Indicates: results of an act or event
time	Indicates: when something happened; refers to a period, point in time, or another event.

mood

A grammatical property, which affects the forms of nouns and verbs.

moods: conditional, imperative, indicative, injunctive, optative, potential, subjunctive. Prohibitive is negative imperative mood.

XXX: affirmative, negative.

Categorization of moods.

- Realis is indicative and generic moods. Realis indicates something is or is not true.
- All others are irrealis.

subjunctive (aka conjunctive mood): expresses a wish, emotion, possibility, judgment, opinion, necessity, or action (that has not yet occurred). Sometimes indistinguishable from the indicative

Montague grammar[1]

morpheme

Basic element of meaning.

Bound morphemes: cannot occur by themselves

Prefixes: attach before other morpheme

Suffixes: attach after other morpheme

Infix: inserted into a morpheme

Some morphemes are not meaningful in isolation but acquire it in combination with others.

Free morpheme: can occur by themselves

Pronunciation: may have several different phonetic forms.

Underlying form: a morpheme which has wider distribution

Allomorph: derived from underlying form by a morphophonemic rule

Plural suffix, past tense suffix,

Derivational morphology. The suffix may change the grammatical category of the underived word

Meaning of derived word is not always predictable

Inflectional morphology. Bounds morphemes are used to indicate grammatical marker. They never change the syntactic category of the words or morphemes to

[1] "Papers in Montague Grammar" Robert Rodman, Linguistics Department, UCLA, Los Angeles, CA 90024 1972

Comments on Richard Montague's "Quantification in Ordinary English" Barbara Partee, in "Approaches to Natural Language"

"Introduction to Montague Semantics" David R Dowty, Robert E Wall, Stanley Peters. Kluwer Academic Publishers (Dordrecht, Netherlands), 1981, printed 1992 Netherlands

"Formal Philosophy: Selected papers of Richard Montague" Edited by Richmond H Thompson, Richard Montague, York University Press. New Haven and London, 1974, Yale University, LCCC 73-77159

which they are attached.

In word order, they usually follow derivational morphemes, and their meaning is predictable.

name Fanciful appellation, nom de plume, nom de guerre, stage name, identity

nouns f :(individual)*

- Collective nouns. Define relation and modal constructions.
- References. Names, pronouns, complexity of query
- Nouns, noun modifiers
- Pronomials – phrase that looks tuff in context, and acts a pronoun.
- Count nouns: rocks, area process
- Mass nouns: water, dough, place, event.
- Abstract nouns

countable nouns[2] Usually treated as an iterator

Some phrase propose a single countable noun. But what if there is more than one? ('the red car' when in fact there are several red cars)

Logicians often treat this the same way as when there are no red cars; this simply is non-sense and wrong.

- Singular countable,
- uncountable,
- plural countable

Noun-Phrase ::= Pre-Modifier Pre-Determiner Determiner Post-Determiner Head-Noun Post-Modifier

grammatical properties Animate: Animate, inanimate

Gender: male, female, neutral

Number: 1,2, more, definite singular, indefinite singular, plural

Case. *See case.* Nouns are typically undeclined, except for possession.

plural nouns Plural countable nouns are one of the simplest: most often they are translated into an iterator.

single nouns Single, countable nouns will be translated into iterators that are converted into scalar

uncountable nouns Most of variables that are set or fetched

Table : *example noun forms*

number		*example*	
singular	indefinite	a beer	
singular	definite	the beer	
singular	definite	all of the beer	the entire portion of a single beer
plural	indefinite	all of the beer	
plural		all of the cars	
plural		all of the mice	collective form

[2]Grimm 2010 Dissertation Proposal

Grimm 2010 Aug, PhD Thesis.
http://parles.upf.edu/llocs/sgrimm/publications/grimm_dissertation.pdf

noun phrases

Nouns are used to reference or predication.

For our purposes there are three types of nouns and noun phrases: uncountable, single, and plural. They are complex in that they are built on top of other complex phrases.

$\{x \mid x \in noun \wedge [\![\alpha]\!](x)\}$

see also *prepositional phrase, predicate*

number agreement

NP_SC = ART_SC N_SC
NP_PL = ART_PC N_PC

nounP ::= art nounP 1
nounP::= nounP pp
nounP ::= noun noun (horse flies)
nounP ::= name
nounP ::= n num
nounP ::= noun
nounP ::= adj nounP

restriction

The prepositional phrase portion may have one of the following roles in the noun phrase:

- argument
- restrictive modifier
- non-restrictive modifier

noun-noun phrases

Olive oil – oil made from olive's
Palm oil – oil made from oil palms
Baby oil – oil for use on babies
Machine oil – oil for use within a machine

objectification

Create a noun from a verb or adjective

one ball bill

Take each word and generate phonetic variants (w/ max distance). Words with the same variant can be used as alliterative:

Ball -> Bill Ba Al : Al Ball
Bill -> Bll Bi ill : Ba Ball
: Bi Bill
: Bl Ball,Bill
: Ill Ill

one-ball-> suffix / prefix string | Markov generation
frame generation

noun-noun → part of speech generation

Onomasty

Study of names, including portions of names (given or *christian* names, family or *surnames*, common or *nicknames*), ordering of name parts and structuring of a full name; variations of transliteration, and the process of mapping this to an individual.

See also *names*

optimality

A strict ranking system

part of speech

Problem with "part of speech" and intuitive interpretation
Time flies like an arrow
Fruit flies like a banana (Grouch Marx)

Almost caught a fish != caught(X,y), fish(Y), Almost(X).

phonetics
acoustic

uses wave theory

possessive case, word form, and punctuation

post modifier

participle verb forms that act as adjectives. Non-finite form of the verb; in English it is used adjectivally and to form compound tenses.

These are typically is a different glossary how to translate to mechanical evaluation.

participle phrase behaves as an adjective modifying a noun or pronoun

::= participle object? modifiers
::= participle complement? modifiers

pragmatics A word or sentence whose reference can't be determined without knowledge of the context of use. Example: I, this, that.

Indexical terms
Egocentric particulars (Russell)
Token reflexive expressions (Reichenbach)
Indicator words (Goodman)
Non-eternal Sentences (Quine)

See also *reference*

predicate Control:

Raising: its subject is not its own argument

prefix Collective prefix, perfective prefix, intensive prefix.

premise indicator as, as indicated by, as shown by, because, follows from, for, for the reason that, in as much as, in view of the fact that, may be deduced from, may be derived from, may be inferred from, since, the reason is that

preposition Not inflected

The *direct object* is to the right; the *object* is to the left.

Rank and space prepositions

$\lambda_{right}\ \lambda_{item}$ [⟦*term*⟧ (right)(item)]

$\lambda_{right}\ \lambda_{item}$ [⟦*term*⟧ (⟦*term**⟧(right)) (⟦*term**⟧(item))]

Where term* is the locator within the concept space; "on" uses place(), rankings use the ranking function for the term

see *modifier*.

as $\lambda_{right}\ \lambda_{p}$ [p(right)]

$\lambda_{right}\ \lambda_{item}$ [right(item)]

of Color of money:

$\lambda_{right}\ \lambda_{p}$ [p(right)]

see *genitive*

spatial preposition The object and direct object may be entities (things) or places; the space may be analytic or topological. In the case of analytic space, entities are mapped to their place in that world at that time and applied to the operation with the space algebra. In the case of topological space, the object and direct object are applied to the relation that is looked in that world at that time.

$\lambda_{right}\ \lambda_{item}$ [⟦*term*⟧ (place(right))(place(item))]

$\lambda_{right}\ \lambda_{item}$ [⟦*term*⟧ (right)(item)]

examples: on, under, over, near,

on, when a graph: ob ∈ On′ (do)

when a topological: before/after

than

$\lambda_{right}\lambda_{op}\,\lambda_{p}$ [op(p(right))]

$\lambda_{right}\lambda_{op}\lambda_{item}$ [op(right(item))]

prepositional phrases

PREPS = PP | PP PREPS;
PP = P | P NP;

template:

$_preposition_phrase:= $_noun_phrase
$_preposition_phrase:= $_noun_phrase $_preposition $_preposition_phrase
$_noun_phrase := $_adj $_noun
$_noun_phrase:= $_nouns

pronoun
indefinite pronoun

any, another, anyone, anything, both (adj), each (adj-sc), either (adj-sc), neither (adj-sc), everybody, few, many, must, much, none (indef adj), several, some (indef adj), such

demonstrative pronoun

Note that a demonstrative can also be used as a determiner

Table : *Demonstratives*

Demonstrative	*number*		*formality*
that	1	distal	standard
this	1	proximal	standard
those	>1	distal	standard
thilk	1		archaic
these	>1	proximal	standard

personal pronoun

also has reflexive form (anaphoric reference)

possessive determiner form

predicate adjectives (a type of possessive)

Table : *Pronoun roles based on case*

Case	*Role*
subjective	subject of a verb
objective	objective of a verb or preposition;
	disjunctive pronoun

Table : *Personal pronouns*

person	*gender*	*number*	*formality*	*subject*	*object*	*reflexive*	*personal pronoun*	*possessive determiner*
1st		singular		I	me	myself	mine	my
1st		plural		we	us	ourselves	ours	our
2nd			standard	you	you	yourself	yours	your
2nd		singular	archaic informal	thou	thee	thyself	thine	thy
2nd		plural	archaic informal	ye	you	yourselves	yours	your
3rd	masculine	singular		he	him	himself	his	his
3rd	feminine	singular		she	her	herself	hers	her
3rd	neuter	singular		it	it	itself	its	its
3rd	generic	singular	formal	one	one	oneself		one's
3rd	generic	singular	nonstandard	they	them	themselves	theirs	their
3rd		plural		they	them	themselves	theirs	their

punctuation

qualifiers [List] [Qualifier]

Qualifier applies to the list antecedent in the list

[List], [Qualifier]

qualifier applies to each item in the list, OTHER than the exceptions to the list. (e.g. a list might be "all animals, except cats")

question answer Considers both the question and sentiment to select the form of the answer.

Set usage guidelines: how questions should be formatted, how answers should be formatted, identification of topic Q&A, frequency of check

Stable responses: review A before final; cataloging Q&A for future use, preventing the same Q from multiple answers

reference

anaphora	When a pronoun refers to a previous noun
cataphora	When a pronoun refers to something later in the text
exophora	When a pronoun refers to something outside of the text

Table : *pronoun references*

sense *confusing, avoid.* Sense typically refers to a distinction in meaning. In idealized languages, it is often axiomatic that all linguistically admissible symbols or expression must have atleast one meaning (sense), and that no two primitive expressions have exactly the same sense. A symbol (or word's) references to a thing, event, etc. is usually considered as being done via sense.

The properties of a sense include: non-psychological, objective, existentially independent, non-empirical, not a property of any mind.

sentiment analysis Approximate model of expressed opinions. Determines:

1. If the opinion is *positive* or *negative*
2. The intensity of the opinion
3. How subjective or impartial the point of view is (eg based on the # of adjectives in the sentence)

Context depends on the topic, and who is expressing themselves, and (to a lesser extent) the forum in which it is expressed. Often the *trend* of sentiment is more important.

uses

1. Used to shape others opinions
2. Used to evaluate effectiveness of a marketing campaign
3. To route in a customer service organization to improve type of service given to that person.

email classification Gather information from email, classify, and respond or route. Key attributes:

- Attitude (negative, neutral, positive). See *sentiment analysis*.
- Issue: none, billing problem, merchandise return, legal..
- Product: individual keywords identify the product or type of product
- Request: eg any of the words "anywhere," "sell," and "me" in the same sentence indicates a request for sales outlet.
- Customer: address, zip code, end consumer, etc.

Find phrases.

Weight their sentiment by the length of the phrase.

	Term is mapped to a positive weight, negative weight, whether it can be neutral.
slang glossary	jargon, slang, obscenities. the origin, characteristics, of the argot. The attitudes and prominent features of the life of those people, and how it is indicated in their speech. Abbreviations, proverbs & sayings, idioms,
speech repairs	(aka disfluencies) when the speaker corrects (or changes) something already said. Heeman and allan
	Fresh starts: Abandons the current utterance and starts again ("So I… luckily we don't have worms this time."
	Modification repairs. Modifies what was said previously. "after the orange juice is at… the oranges are at the OJ factory"
	Abridge repairs. Consists solely of a fragment and/or editing term (e.g. pauses) "we had a line, but *ah* I think…"
spell checker	Take a dictionary of properly spelled words, and misspell each in a variety of ways. Map each misspelling back to the proper word, with a Levenshtein distance – typically retaining only one proper spelling for a given misspelling (the one with lowest distance). Use this mapping to offer correction suggestions.
predictive	lexicon stored as a trie, on a per letter basis, looks up estimates (frequency of) next character.
suprasegmentals	stress, tone, syllabification. Often these are what people mean by "accent" These get lost in singing;
syntax	
center embedding	A is nested in B with non-theta materials to the left and right. Cat(A) == Cat(B)
closure issues	Humans use 2 strategies: minimal attachment – early closure (Kimball: early closure, attach high if possible). Right associative – late closure attach low if possible.
left branching	A is nest in B with non-theta material to the right, Cat(A) == Cat(B).
right branching	A is nested in B with non-theta materials on the left, Cat(A) == Cat(B)
text analysis	Sentiment analysis
	Information extraction, Topic tracking, summarization, categorization, clustering, concept linkage, how questions are answered
	see also *transcribe, translate, transliterate*
information extraction	identifying key phrases, identifying key relationships. Predefined sequences or patterns
summarization	reduction in length and detail, but retains main point & key details. Based on estimate of semantic weight, position, patterns
categorization	groups documents by main theme (predefined set usually), identifies main theme without information extraction.
clustering	groups documents
topics	Types of topics: names of things, authors, people, places, event, time, other nouns
	types of occurrence: definition, synopsis, see also associated topics
	topics have: name, occurrence, rules in association, relationship with other topics (has-a, is-a, arcs – verbs, relationship), rules: had better, must, enforce, concept of subject identity.
	Document elements: thesis, reason, anti, example, deny
transcribe	write down what is said

translation concerned with preserving meaning and intent. It depends on:

- subject area
- source language
- target language

series of events; catalog of choices

translation of certain construct

analysis preliminary analysis:

form & context of the text as a whole

semantic & stylistic analysis: constructs, idioms, style of speech

levels levels:

1. message and content
2. syntax
3. lexicon

transliteration Transcribe combination of letters from one orthography or alphabet to another. include ideograph

types of writing informative writing: consists of more nouns, adjectives, prepositions, determiners, and coordinating conjunctions

imaginative writing: more verbs, adverbs, pronouns, pre-determiners

deceptive writing: similar to imaginative, except more adjectives and adverb superlatives

Rayson 2001

P Rayson, A Wilson, G Leech 2001, Grammatical word class variation within the British National Corpus sampler. *Language and Compilers, 36(1):295-306*

Mye Ott, Yejin Choi, clair Cardie, Jeffrey Hancock "Finding Deceptive Opinion Spam by Any Strectch of the Imagination" *ACL HLT 2011*

vagueness A word's lack of precision.

see also *modifier (dangling and vague)*

quantitative vagueness A word may be considered unacceptably vague when there is a need or wish to replace the word with a quantitatively more precise expression – the required level of precision is higher than that provided.

example Quantifers: We bought fresh peaches, apples, cherries, and pears, at the orchard. Atleast the peaches were fresh. People don't seem bothered about which fruits were fresh.

task-related A word is considered vague when:

We need to decide whether the word applies in a particular case

We are uncertain about whether the word applies

We cannot resolve the doubt by acquiring additional facts

unacceptably vague When we must accomplish a certain task and are blocked from doing so by doubt about how the word is to be applied to actual cases.

studiedly vague Language that is deliberately vague, and

Certain crucial words and phrases have been carefully chosen so as not offend the doctrines and principles of those who must approve the document.. (eg legal documents or agreements often prefer an open texture)

verb grammatical attributes Voice: active, passive

Person

tense

Mood: see *mood*

aspect: ongoing or done

see also *modifier*

linking verb — subject to noun or modifier

mood — Represents modality (possibility, etc). Whether the verb is a statement of fact or is about possibility.

Finite forms, conditional, imperative, indicative, injunctive, optative, potential, subjunctive. Infinite forms, gerund forms, participles.

aktionsart — A classification of verbs into ~five classes based on their temporal aspect

States: know, love, be beautiful, be on time

activities: run push, draw

Accomplishments: cross the street, write a letter

Achievements: being, reach, arrive

Points: flash, spot, blink

verb phrase

vp ::= v
vp ::= v np
vp ::= verb pp
vp ::= aux vp
vp ::= ModalAuxVerb VP
vp ::= ModalAuxVerb negation VP
vp := verb np pp
vp ::= v inf
inf ::= « to » vp

verbal auxiliary — Helper verb, e.g. to be and to have.

to be

$\lambda_{right}\ \lambda_{left}$ [right == left]

$\lambda_{right}\ \lambda_{left}$ [right(left)]

$\lambda_{right}\ \lambda_{left}\{e|\ e \in left \wedge right(e)]\}$

to have

$\lambda_{right}\ \lambda_{left}$ [right(left)]

$\lambda_{right}\ \lambda_{left}$ [right(possess(left))]

vowel harmony — "Vowel harmony is a process that results in all vowels of a word sharing a certain feature or features. Morphophonemic rules of vowel harmony are found in many languages." *Ling200 class handout*

word division — see *hyphenation*

word probability — A function of counts.

See also *Stop lists, heap's law, Zipf's law*

Noun prepositions: of

Zipf's law — Some words are disproportionately used. The frequency of words used in a language follows a distribution:

$$= \frac{1}{r} \log w$$

w= the number of words in the language

r = rank or order in the list