1 Fundamental Concepts

Reference: How does the mapping between form and meaning work? Does it work at all?
Compositionality: How are complex utterances built from smaller units? Are they built from smaller units at all?
Combinatorial structure: a small number of meaningless building blocks (phonemes, parts of syllables) combined into an unlimited set of utterances (words and morphemes)
Compositional structure: meaningful building blocks (words and morphemes) are combined into larger meaningful utterances (phrases and sentences)

2 Information Theory

Information Content (Surprisal)
$-log_2 p(x)$: the more frequent the word, the lower its information content. e.g. the word type "blue" occurs ca. 3750 times in 10000 tokens, and its information content is $-log_2(3750/10000) \approx 1.42$ bits.

Shannon Entropy
$H(X) = -\sum p(x) log_2 p(x)$: entropy as probability, the average information content of information encoding units in the language. Measure of information encoding potential of a symbol system. The higher the uncertainty, the larger the entropy. e.g. $H_{char}(Morse) = -(\frac{86}{136} * log_2(\frac{86}{136}) + \frac{50}{136} * log_2(\frac{50}{136})) \approx 0.949$ bits per character.
A series of studies proposed to use entropic measures to distinguish human writing from other types of symbol systems.

Joint Entropy, Conditional Entropy
Joint Entropy: $H(X,Y) = -\sum\sum p(x,y) log_2 p(x,y)$
Conditional Entropy: $H(Y|X) = -\sum p(x) \sum p(y|x) log_2 p(y|x)$ The more ambiguity in language (uncertainty), the higher conditional entropy. No ambiguity -> 0.
Ackerman & Malouf (2013) propose two entropic measures for morphological complexity: the average entropy of a paradigm as a measure of enumerative complexity, and the average conditional entropy of cells as an integrative complexity measure. They are related to learnability.

Probability Estimation

Maximum Likelihood (ML)
Problems: unit problem, sample size problem, interdependence problem, extrapolation problem
Methods: frequency-based, language models, experiments with humans

Mutual Information

$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$ reduction in the uncertainty of X given Y.
compromise between minimum learning cost H(Y) and maximum expliciteness I(X; Y).
Entropy is the upper bound on the mutual information between forms and meanings
Is the entropy rate zero? -> asymptotic determinism of human utterances.

3 Propositional Logic

Why formal logic? overcome ambiguity, determine relationships between meanings of sentences, determine meanings of setences, model compositionality, recursive system.

Definition

Proposition: The meaning of a simple declarative sentence
Extensions: real-world situations they refer to.
Frege's Generalization: The extension of a sentence S is its truth value
The proposition expressed by a sentence is the set of possible cases [situations] of which that sentence is true.
propositional variables: p, q, r...
propositional operators: $\neg, conjunction \wedge, disjunction \vee, XOR, \rightarrow, \leftrightarrow$

Syntax: Recursive Definition

(i) Propositional letters in the vocabulary of L are formulas in L.
(ii) If ϕ is a formula in L, then $\neg\phi$ is too.
(iii) If ϕ and ψ are formulas in L, then $(\phi \wedge \psi), (\phi \vee \psi), (\phi \rightarrow \psi), (\phi \leftrightarrow \psi)$ are too.
(iv) Only that which can be generated by the clauses (i)-(iii) in a finite number of steps is a formula in L.
invalid: $\neg(\neg\neg p), \neg((p \wedge q))$

Construction Trees

- $(\neg(p \vee q) \rightarrow \neg\neg q) \leftrightarrow r$ $(iii. \leftrightarrow)$
  - $(\neg(p \vee q) \rightarrow \neg\neg q)$ $(iii. \rightarrow)$
    - $\neg(p \vee q)$ (ii)
      - $p \vee q$ $(iii.\vee)$
        - p (i)
        - q (i)
    - $\neg\neg q$ (ii)
      - $\neg q$ (ii)
        - q (i)
  - $r(i)$

Valuation Functions

For every valuation V and for all formulas ϕ: $V(\phi \leftrightarrow \psi) = 1 iff V(\phi) = V(\psi)$.

4 Predicate Logic

Introduce constants and variables representing invididuals and predicates to capture the main structural building blocks of sentences.
Introduce quantifiers to allow for quantified statements.

Definition

constant symbols: a, b, c, ...
variable symbols: x, y, z, ...
n-ary/n-place predicate symbols: A, B, C, reflect relations between n elements (n>0)
connectives: $\neg, \wedge, \vee, \rightarrow, ...$
quantifiers: $\forall, \exists$
round brackets (), equal sign =

Syntax: Recursive Definition

(i) If A is an n-ary predicate letter in the vocabulary of L, and each of t1,..., tn is a constant or a variable in the vocabulary of L, then At1,..., tn is a formula in L.
(ii) If ϕ is a formula in L, then $\neg\phi$ is too.
(iii) If ϕ and ψ are formulas in L, then $(\phi \wedge \psi), (\phi \vee \psi), (\phi \rightarrow \psi), (\phi \leftrightarrow \psi)$ are too.
(iv) If ϕ is a formula in L and x is a variable, then $\forall x\phi$ and $\exists x\phi$ is too. (v) Only that which can be generated by the clauses (i)-(iv) in a finite number of steps is a formula in L.
invalid: $a, A, \forall(Axy)$

Construction Trees

- $\forall_x \forall_y((A_{xy} \wedge B_y) \rightarrow \exists_x A_{xb})$ $(iv.\forall)$
  - $\forall_y((A_{xy} \wedge B_y) \rightarrow \exists_x A_{xb})$ $(iv.\forall)$
    - $(A_{xy} \wedge B_y) \rightarrow \exists_x A_{xb}$ $(iii. \rightarrow)$
      - $A_{xy} \wedge B_y$ $(iii.\wedge)$
        - A_{xy} (i)
        - B_y (i)
      - $\exists_x A_{xb}$ $(iv.\exists)$
        - A_{xb} (i)

Semantics
domain (D): set of entities
interpretation functions I = $\{< m, e >, < s, e >, < v, e >\}$, I(m) = e, I(s) = e, I(v) = e.
model M: consists of a domain D and an interpretation function I which conforms to: (i) if c is a constant in L, then I(c) ∈ D
(ii) if B is an n-ary prpedicate letter in L, then I(B) ⊂ D
valuation function V_M:
If Aa1,...,an is an atomic sentence in L, then V_M(Aa1,...,an) = 1 iff <I(a1),...,I(an)> ∈ I(A).
...
$V_M(\forall x\phi) = 1$ iff $V_M([c/x]\phi) = 1$ for all constants c in L.
$V_M(\exists x\phi) = 1$ iff $V_M([c/x]\phi) = 1$ for at least one constant c in L.
If $V_M(\phi) = 1$, then ϕ is said to be true in model M.

Formula vs. Sentence
A sentence is a formula in L which lacks free variables.
Sentence: Aa, $\forall x(Fx)$, $\forall x(Ax \rightarrow \exists y By)$
Not a sentence (but Formula): Ax, Fx, $Ax \rightarrow \exists y By$

5 Second-Order Logic

e.g. CR (CX: X is a predicate with the property of being a color; Rx: x is red)
$\exists X(CX \wedge Xm)$: Mars has a color.
$\exists X(Xj \wedge Xp)$: John has at least one thing in common with Peter.
$\forall x(\exists X((AX \wedge Xx) \wedge Jx) \rightarrow \exists Y(Yx \wedge CY))$: All animals that live in the jungle have a color.
$\forall x(\exists X((Xx \wedge AX) \wedge Gx) \rightarrow Ex)$: If an animal is grey, then it is an elephant.

Vocabulary extention
1st-order predicate variables: X, Y, Z
2nd-order predicate constants: A, B, C
e.g. AX: X is a property with the property of being an animal

Syntax: Recursive Definition
(i) If A is an n-ary first-order predicate letter in the vocabulary of L, and each of t1,..., tn are individual terms in L, then At1,..., tn is an (atomic) formula in L.
(ii) If X is a [first-order] predicate variable and t is an individual term (both constants and variables) in L, then Xt is an atomic formula in L;
(iii) If A is an n-ary second-order predicate letter/constant in L, and T1,...,Tn are first-order **unary** predicate constants, or predicate variables, in L, then AT1,...,Tn is an (atomic) formula in L;
(iv) If ϕ is a formula in L, then $\neg\phi$ is too.
(v) If ϕ and ψ are formulas in L, then $(\phi \wedge \psi), (\phi \vee \psi), (\phi \rightarrow \psi), (\phi \leftrightarrow \psi)$ are too.
(vii) If ϕ is a formula in L and x is a variable, then $\forall x\phi$ and $\exists x\phi$ is too.
(viii) If X is a [first-order] predicate variable, and ϕ is a formula in L, then $\forall X\phi$ and $\exists X\phi$ is too.
(viiii) Only that which can be generated by the clauses (i)-(vii) in a finite number of steps is a formula in L.
invalid: $x, X, Xab, \forall(Xa)$

Semantics
just as a 1st-order predicate denotes a set of entities, a 2nd-order predicate denotes a set of a set of entities.

6 Type theory

Tools to get to grips with frequent compositional structures in natural language (adj-n, adv-v, art-n, prep-np... combis), a higher-order logic

Definition
(i) $e, t \in T$
(ii) if $a, b \in T$, then $< a, b > \in T$
(iii) nothing is an element of T except on the basis of clauses (i) and (ii).
invalid: et, <e,e,t>, <e,<e,t>

Functional Application
If α = <e,t> and β = e then $\alpha(\beta)$ = t.
If α = <t,<t,e» and β = <t,e> then $\alpha(\beta)$ is not defined.

Semantic Types
individual: e
sentences: t
1-place predicates (intransitive verb): <e,t>
2-place predicates (transitive verb): <e,<e,t»
3-place predicates (ditransitive verb): <e,<e,<e,t»>
common nouns (e.g. dog): <e,t>
NP (e.g. the dog): e
determiners (e.g. the): <e,<e,t»
adjectives: <e,t>
adjectives as predicate modifiers(e.g. *happy* dog): «e,t>,<e,t»
adverbs(predicate modifier): «e,t>,<e,t»
sentence modifier(e.g. not): <t,t>
function(entity to entity) (e.g. the farther of): <e, e>
1-place 2nd-order predicate: «e,t>,t>
2-place 2nd-order predicate: «e,t>,«e,t>,t»

Syntax: Recursive Definition
(i) If α is a variable or a constant of type a in L, then α is an expression of type a in L.
(ii) If α is an expression of type <a,b> in L, and β is an expression of type a in L, then $(\alpha(\beta))$ is an expression of type b in L.
(iii) If ϕ and ψ are formulas in L, then so are $\neg\phi, (\phi \wedge \psi), (\phi \vee \psi), (\phi \rightarrow \psi), (\phi \leftrightarrow \psi)$.
(iv) If ϕ is an expression *of type t* in L and v is a variable (of arbitrary type a), then $\forall v\phi$ and $\exists v\phi$ are expression of type t in L.
(v) If α and $beta$ are expressions in L which belong to the same type, then $(\alpha = \beta)$ is an expression of type t in L.
(vi) Every expression L is to be constructed by means of (i)-(v) in a finite number of steps.
formulas: expressions of type t.
Difference to Predicate Logic:
Jumbo befriends Maya.
Predicate Logic: Bjm
Type-theoretic logic: (B(m))(j) or alternatively B(m)(j)

Semantics
Truth valuations via paticular interpretation functions defined for different types of expressions.
e.g. interpretation function I for which it holds that: I(W)(d) = 1 iff d∈W, otherwise 0.

7 Lambda Calculus

To represent parts of sentences or predicates in a fully compositional account. Allows to capture the compositionality of language

Syntax
Add another clause to the type-theoretic language syntax
(vii) If α is an expression of type a in L, and v is a variable of type

b, then $\lambda v(\alpha)$ is an expression of type <b, a> in L.

Lambda-Abstraction

We say that $\lambda v(\alpha)$ has been formed from α by abstraction over the formerly free variable v. Hence, the free occurrences of v in α are now bound by the λ-operator λx.

e.g. expression: S(x) of type t

λ-abstraction: λx(S(x)) of type <e,t>

λx(x) of type <e,e>

λx(B(y)(x)) of type <e,t>

λX(X(a) $\wedge$ X(b)) of type «e,t>,t>

Lambda-Conversion

remove the λ-operator and plug an expression into every occurrence of the variable which is bound by the λ-operator.

e.g. λx(S(x))(c) = S(c)

λx(λy(A(y)(x)))(c)(d)=λy(A(y)(c))(d)=A(d)(c)

λ-conversion is only valid when vairable v is not bound by a quantifier $\forall$ or $\exists$

Modelling Compositionality

John smokes: λx(S(x))(j)=S(j)

smokes: λx(S(x))

smokes and drinks: λx(S(x)$\wedge$D(x))

Jumbo is grey: λx(G(x))(j)=G(j)

is grey: λx(G(x))

Jumbo is: λX(X(j)

is: λX(λx(X(x)))

Truth Valuation

For all entities d in the domain D it holds that h(d)=1 iff I(W)(d)=1. This illustrates that the denotation if $\lambda x(W(x))$ is indeed the same as one would expect for just the word walks represented by W.

8 Modality

A category of linguistic meaning having to do with the expression of possibility and necessity.

Modal Strength(Force)

Statements can express stronger or weaker commitment to the truth of base proposition.

High: Arthur must/has to be home.

Medium: Arthur should be home.

Low: Arthur might/could be home.

Modal Type(Flavor)

Epistemic modality: relative to speaker's knowledge of the situation

John didn't show up for work. He must be sick. (spoken by coworker)

The older students might/may(?) leave school early (unless the teachers watch them carefully).

It has to be raining. [Seeing people outside with umbrellas]

Deontic: relative to authoritative person or code of conduct

John didn't show up for work. He must be sick. (spoken by boss)

Visitors have to leave by 6pm.

Dynamic: Concerned with properties and dispositions of persons

John has to sneeze.

Anne is very strong. She can list this table.

Teleological: achieving goals or serving a purpose

To get home in time, you have to take a taxi.

Anne must be in Paris at 5pm. She can/must take the train to go there.

Polysemy Controversy

In some languages, modal auxiliaries can be used for different types of modality.

Ambiguity(polysemy) vs. Indeterminacy

Contradiction test (If a sentence of the form *X but not X* can be true, then expression must be ambiguous.

e.g. They are not children any more, but they are still my children.

John must be sick, but he must not be sick.

John can be sick, but he cannot be sick...

If considered non-contradictory, then the modal auxiliaries are polysemous with regards to modal type

Adverbial Phrase Test:

e.g. Dynamic: (In view of his physical abilities,) John can lift 200 kg.

If adverbial phrases in parentheses are not redundant, type of modality is not lexically specidied but inferred from context, i.e. indeterminate

Modal Logical Operators

$\Diamond p$: it is possible that p...

$\Box p$: it is necessary that p...

modality as quantification over possible worlds: $\Diamond p \equiv \exists w[w \in p]$, $\Box p \equiv \forall w[w \in p]$

Modal propositional logic: add one more syntactic clause to the syntax of propositional logic:

(v) if ϕ is a formula in L, then $\Box\phi$ and $\Diamond\phi$ are too.

valid formulas: $\Box\Diamond p$, $\neg\Diamond(p \wedge q)$, $p \rightarrow \Box\Diamond p$

Fundamental tautologies:

$\Diamond\phi \leftrightarrow \neg\Box\neg\phi$: something is possible if and only if it is not the case that it is necessarily not the case

$\Box\phi \leftrightarrow \neg\Diamond\neg\phi$: something is necessary if and only if it is not the case that it is possibly not the case.

Modality and Truth-Conditions

Both epistemic and root modality can be part of the proposition and contribute to its truth conditions.

Challenge Test: Is the epistemic modal marker part of what can be challenged about a proposition?

A: John profited from the old man's death, he must be the murderer. B: That's not true; he could be the murderer, but he doesn't have to be.

Yes-No Question Test: Can the epistemic modal marker be the focus of a yes-no question?

A: Must John be the murderer? B: Yes, he must. or: No, he doesn't have to be.

Negation Test: does negation scope over and hence include the modal marker as part of the negated proposition?

Smith cannot be the candidate. $\neg\Diamond p$ √

Smith might not be the candidate. $\Diamond\neg p$

8.1 Cross-Linguistic Variation

Epistemic possibility: verbal constructions/affixes on verbs/other

Situational possibility: affixes on verbs/verbal constructions/other

9 Evidentiality

covers the way in which information was acquired, without necessarily relating to the degree of speaker's certainty concerning the statement or whether it is true or not. To be considered as an evidential, a morpheme has to have 'source of information' as its core meaning; that is, the unmarked, or default interpretation.

Definition

1st claim: It is a "linguistic category", i.e. a grammatical category with grammatical markers (same as for modality).

2nd claim: These evidential markers have source of information as their core meaning.

markers can develop polysemy, e.g. tense marking and evidential marking, can be used recursively without being redundant

3rd claim: Evidentiality is not "necessarily relating to the degree of speaker's certainty", i.e. it is distinct from epistemic modality.

Evidentiality vs. Epistemic Modality

There is good evidence that evidential markers in a number of languages do not contribute to propositional content but function as illocutionary modifiers, and so must be distinct from epistemic modality.

Negation Test: If negation can scope over the evidential marker, then the evidential marker is considered to contribute to the truth-conditional content.

Challenge Test: The hearer can challenge the truth of the statement of the speaker given more direct evidence, but the source of information cannot be challenged.

Two types of evidentials:

Illocutionary: markers of evidentiality that do not contribute to the truth-conditional content, but that "add to or modify the sincerity conditions of the [speech] act"

Propositional: markers of evidentiality that also contribute to the truth-conditional content. e.g. Es soll regnen.

9.1 Cross-Linguistic Variation

Semantic distinctions of evidentiality: no grammatical evidentials/indirect only/direct and indirect

Coding of Evidentiality: no grammatical evidentials/verbal affix or clitic/part of the tense system/separate particle/modal morpheme/mixed

10 Introduction to Pragmatics

Semantics: word meaning, sentence meaning Pragmatics: utterance meaning

Definitions

Anomaly Definition:study of those principles that will account for why a certain set of sentences are anomalous, or not possible utterances.

Functional: attempts to explain facets of linguistic structure by reference to non-linguistic pressures and causes.

Context: part of performance, explicate the reasoning of speakers and hearers in working out the correlation in a context of a sentence token with a proposition.

Grammaticalization: study of those relations between language and context that are grammaticalized, or encoded in the structure of a language.

Truth-Conditional: those aspects of the meaning of utterances which cannot be accounted for by straightforward reference to the truth conditions of the sentences uttered.

Inter-Relation: interation of context-dependent aspects of language structure and principles of language usage, relations between language and context

Appropriateness/Felicity: study of the ability of language users to pair sentences with the contexts in which they would be appropriate.

List: study of deixis (at least in part), implicature, presupposition, speech acts, and aspects of discourse structure.

More promising: Inter-Relation, Truth-Conditional

11 Discourse Representation Theory

To deal with issues in the semantics and pragmatics of anaphora and tense

Discourse representation structures:a hearer builds up a mental representation of the discourse as it unfolds, and that every incoming sentence prompts additions to that representation.

Anaphora Resolution

Anaphora as co-reference: *John* likes *his* donkey.

Anaphora as binding: *No farmer* likes *his* donkey.

Anaphora as neither co-reference nor binding: John owns *a donkey*. *It* is grey.

Discourse Representation Structures

Merging: [x, y: farmer(x), donkey(y), chased(x,y)] + [v, w: caught(v, w)] = [x, y, v, w: farmer(x), donkey(y), chased(x,y), caught(v, w)]

Anaphora Resolution: = [x, y, v, w: v=x, w=y, farmer(x), donkey(y), chased(x,y), caught(v,w)] = [x, y: farmer(x), donkey(y), chased(x,y), caught(x,y)]

Complex DRS Conditions

Negation: John doesn't own a donkey. It is grey. [1x,z: John(x), ¬[2y: donkey(y), owns(x,y)], grey(z)]. y is not accessible to z. x and z are accessible to y.

Conditionals: If John owns a donkey, he likes it. [1:[2x,y: John(x), donkey(y), owns(x,y)]→[3v,w: likes(v,w)]]=[1:[2x,y,v,w: v=x, w=y, John(x), donkey(y), owns(x,y)]→[3: likes(v,w)]]=[1:[2x,y: John(x), donkey(y), owns(x,y)]→[3: likes(x,y)]]. x and y are accessible to v and w.

Quantification: Every farmer who owns a donkey, likes it. [1:[2x,y: farmer(x), donkey(y), owns(x,y)]$\forall x$[3v,w: likes(v,w)]]