Processing pragmatic and referential information I
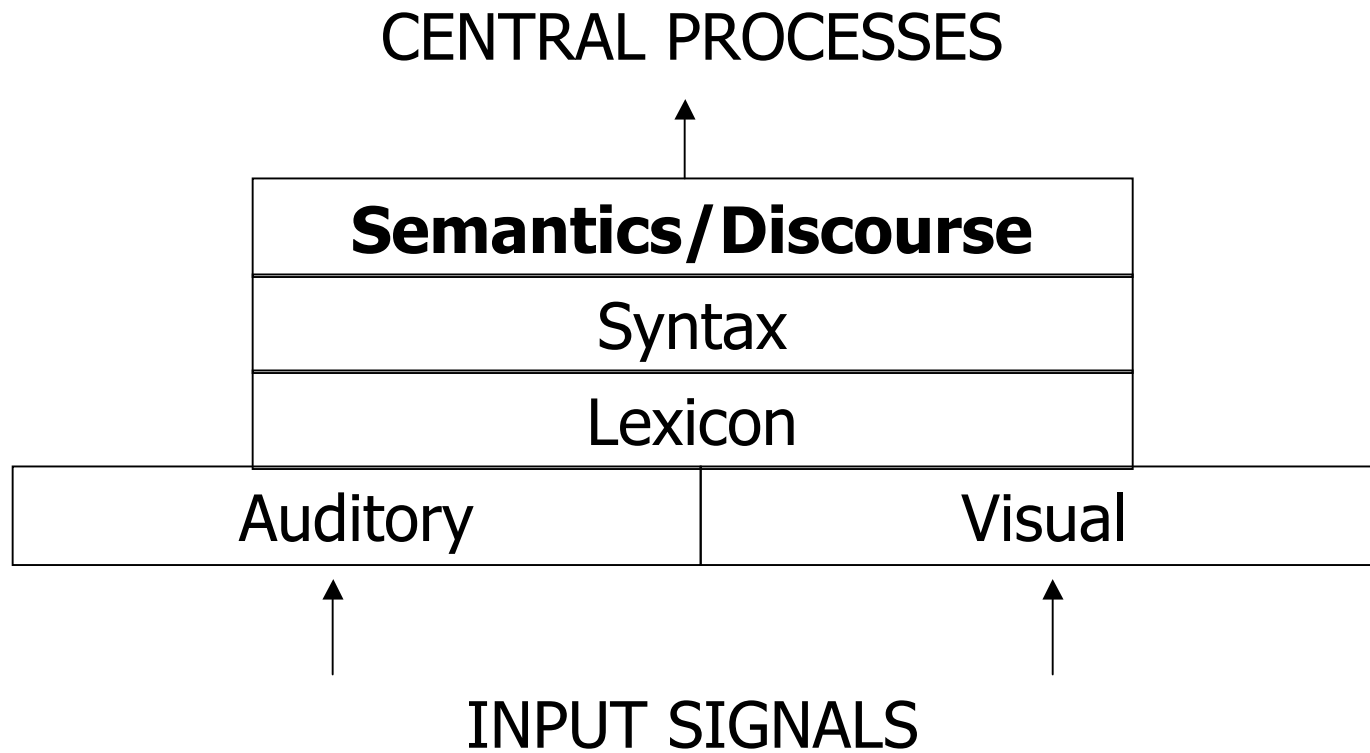
9.59 / 24.905

March 8, 2005

Ted Gibson

On-Line Language Processor

CENTRAL PROCESSES

| **Semantics/Discourse** |
| Syntax |
| Lexicon |

| Auditory | Visual |

INPUT SIGNALS

To do today and next time: non-syntactic information in sentence comprehension

1. Information flow in sentence processing
   - Grammatical constructions have **functions.** E.g., Restrictive modifier: pick an item from a set.

2. Figurative / non-literal language:
   1. Common types of figurative language
   2. Interpreting figurative language
   3. Grice's Maxims of cooperative conversation
      - Implicature: the non-literal information that an utterance conveys

3. One theory of processing some non-literal language: **The Referential Theory**
   - The principle of parsimony
   - The principle of referential support

4. Modularity re-visited: Can discourse information **guide** syntactic processing?

5. Pragmatic processing: Implicatures and on-line sentence processing
   - When do people compute contrast sets?

Information flow and sentence complexity

Edward Gibson, Timothy Desmet, Dan Grodner, Duane Watson & Kara Ko

Nested structures are more complex than non-nested structures

(1) # The student who the professor who the scientist collaborated with had advised copied the article.

(2) The scientist collaborated with the professor who had advised the student who copied the article.

The Dependency Locality Theory (DLT, Gibson, 1998, 2000)

Resources are required for two aspects of language comprehension:

1. Integration: connecting the current word into the structure built thus far;

2. Storage: Predicting categories to complete the current structure.

   - **Syntactic predictions**: processing cost for each head that is required to complete the current string as a grammatical sentence

2 Predictions of storage cost theory:

Prediction 1: People should read (1) faster than the same region in (2):

(1) **The reporter who the senator attacked ignored the president**.

(2) The fact that **the reporter who the senator attacked ignored the president** bothered the editor.

Prediction 2:

Object-modifying RCs as in (2) should be processed faster than subject-modifying RCs as in (1):

(1) The reporter **who the senator attacked** ignored the president.

(2) The president ignored the reporter **who the senator attacked.**

All theories of nesting make this prediction.

Experiment 1

Three factors:

(1) Modifier position: subject-modifying, object-modifying

(2) Extraction type: subject-extracted, object-extracted

(3) Embedding: Embedded, not embedded

Experiment 1 materials

(a) Subject modifier, object-extracted (SO), not embedded
The reporter **who the senator attacked** on Tuesday ignored the president.

(b) Object modifier, object-extracted (OO), not embedded
The president ignored the reporter **who the senator attacked** on Tuesday.

(c) Subject modifier, subject-extracted (SS), not embedded
The reporter **who attacked the senator** on Tuesday ignored the president.

(d) Object modifier, subject-extracted (OS), not embedded
The president ignored the reporter **who attacked the senator** on Tuesday.

Experiment 1 materials (continued)

(e) Subject modifier, object-extracted (SO), embedded

The fact that the reporter **who the senator attacked** on Tuesday ignored the president bothered the editor.

(f) Object modifier, object-extracted (OO), embedded

The fact that the president ignored the reporter **who the senator attacked** on Tuesday bothered the editor.

(g) Subject modifier, subject-extracted (SS), embedded

The fact that the reporter **who attacked the senator** on Tuesday ignored the president bothered the editor.

(h) Object modifier, subject-extracted (OS), embedded

The fact that the president ignored the reporter **who attacked the senator** on Tuesday bothered the editor.

Experiment 1 predictions

(1) Subject-extractions will be processed faster than object-extractions

(2) Object-modifiers will be processed faster than subject-modifiers

(3) Unembedded RCs will be processed faster than embedded RCs

Experiment 1 residual reading times in the RC **who the senator attacked / attacked the senator**, as a function of modifier type, extraction type and embedding
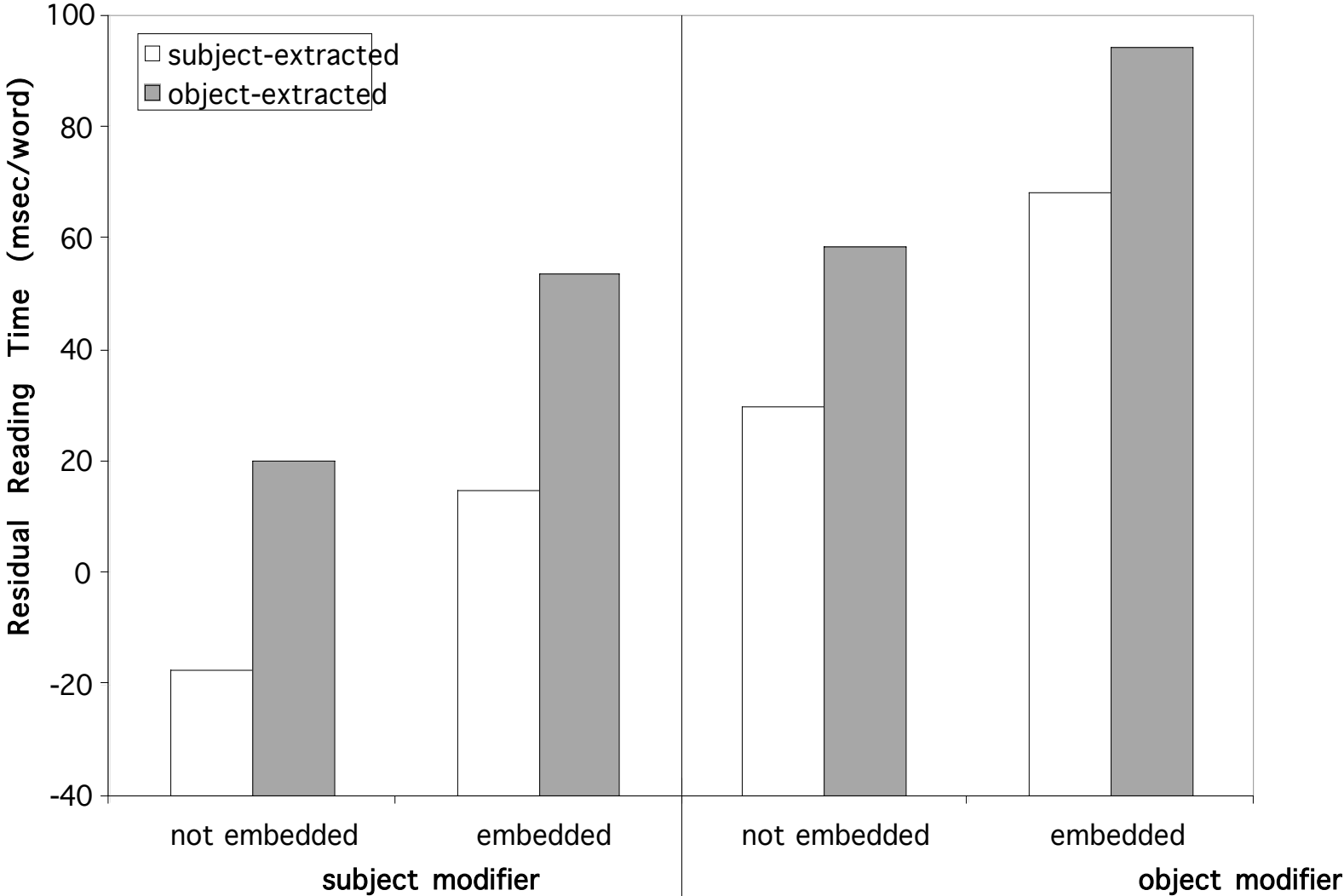
Three main effects:

- Correct prediction of integration costs: subject-extractions are processed faster than object-extractions

- Correct prediction of storage costs: unembedded RCs are processed faster than embedded RCs

- **Incorrect** prediction of storage costs: subject-modifiers are processed **faster** than object-modifiers

How to account for the final result? This result is totally unpredicted by the DLT and all other resource theories.

What is a relative clause?

There are at least two kinds of RCs:

(1) The student that studied for the exam aced the test.
Restrictive RC: no pauses separating RCs

(2) Mary, who studied for the exam, aced the test.
Non-restrictive RC: pauses separating RCs

What is a relative clause?

Restrictive modifiers (e.g., (1)) usually serve a contrastive function: picking an element from a set.

I.e., They identify a particular referent from among a group of entities that contrast along the dimension denoted by the modifier.

E.g., "the student that studied for the exam" is usually used when there is a set of students out of which one can be identified by the restrictive modifier.

Consequently, the information in a restrictive RC is usually **background** information.

(More about non-restrictive RCs in a moment…)

The information flow hypothesis

Old, background information is comprehended more easily early in a sentence, such as in a position modifying the subject; New, foreground material is processed more easily later in a sentence, such as in a position in the main predicate of the sentence.

(Chafe, 1976, 1987; Du Bois, 1987; Givon, 1979, 1983, 1984; Prince, 1981)

By the information flow hypothesis, restrictive RCs should therefore be processed faster at the beginning of a sentence, where background information usually appears.

Hence, according to the information flow hypothesis, subject-modifying RCs will be processed faster that object-modifying RCs, as observed in Experiment 1.

Non-restrictive modifiers

(2) Mary, who studied for the exam, aced the test.

Non-restrictive modifiers add extra "aside" information. They do **not** function as identifying information in a set denoted by the head noun:

(3) My father, who ate ham this morning, became extremely ill.

(4) The sun, which rises in the east, can be used to orient oneself.

Non-restrictive modifiers

Cues to a restrictive RC in English: no pauses (commas) around the RC; the use of the complementizer "that"

Cues to a non-restrictive RC in English: pauses (commas) around the RC; the use of wh-pronouns "who" or "which"

(3) My father, who ate ham this morning, became extremely ill.

(5) * My father that ate ham this morning became extremely ill.

(4) The sun, which rises in the east, can be used to orient oneself.

(6) * The sun that rises in the east can be used to orient oneself.

Prediction:

Non-restrictive RCs, which introduce some new material, and are not typically background material (not really foreground either: more like asides), should be processed differently than restrictive RCs.

Experiment 2 materials:

(4) (a) Subject-modifier, restrictive

A group of film critics praised a director at a banquet and another director at a film premiere. The director **that the critics praised at a banquet** insulted an actor from a big action movie during an interview.

(b) Object-modifier, restrictive

A group of film critics praised a director at a banquet and another director at a film premiere. An actor from a big action movie insulted the director **that the critics praised at a banquet** during an interview.
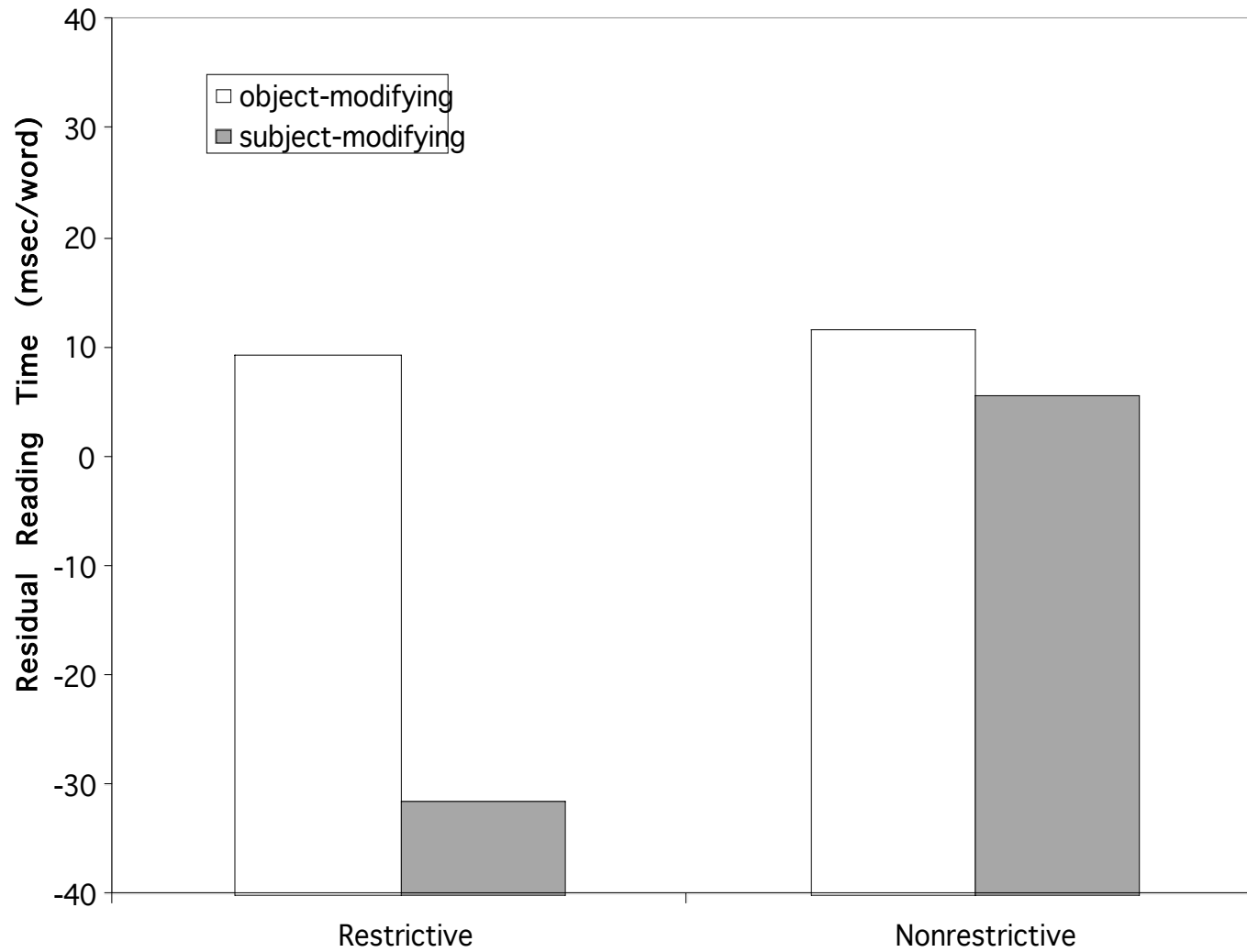
(c) Subject-modifier, non-restrictive

A group of film critics praised a director and a producer. The director, **who the critics praised at a banquet**, insulted an actor from a big action movie during an interview.

(d) Object-modifier, non-restrictive

A group of film critics praised a director and a producer. An actor from a big action movie insulted the director, **who the critics praised at a banquet**, during an interview.

Experiment 2 residual reading times in the RC, as a function of modifier type and restrictiveness.
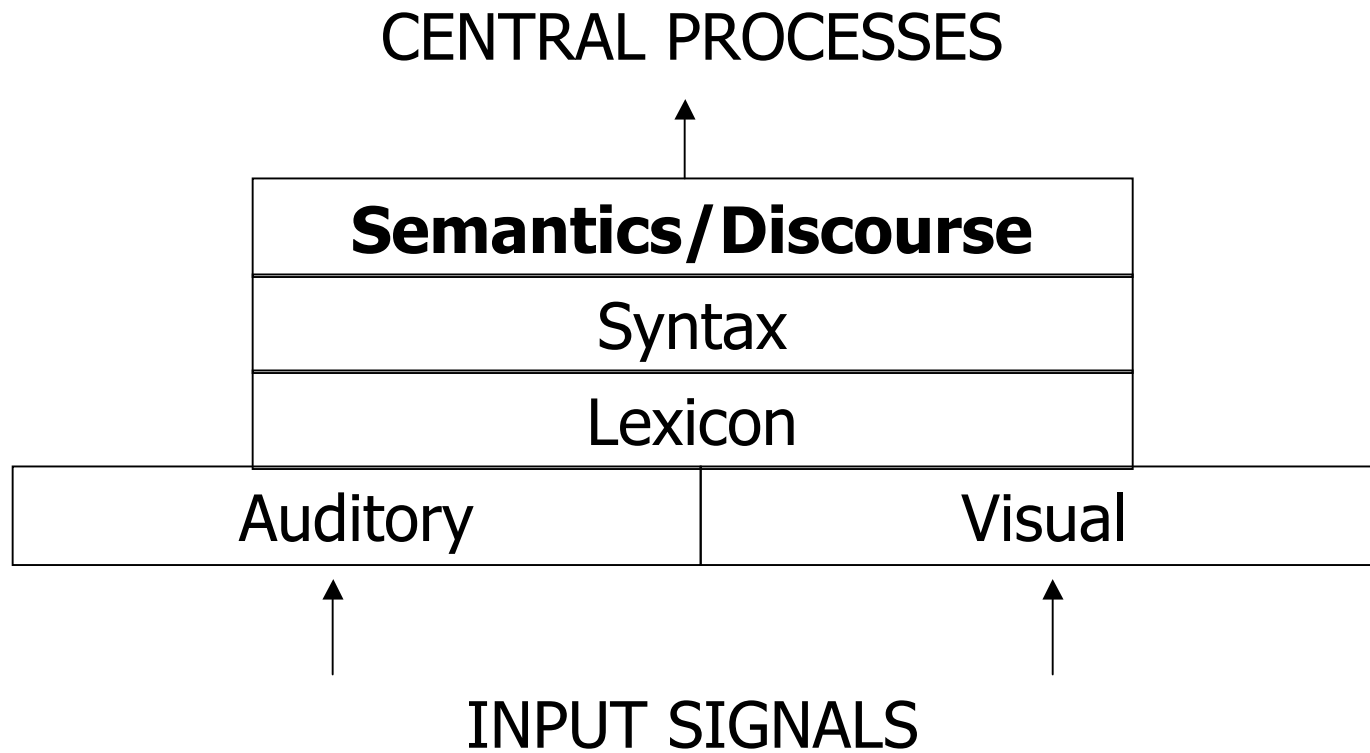
Results

- Replication of Experiment 1: subject-modifying restrictive RCs are read faster than object-modifying restrictive RCs

- Non-restrictives do not show the same advantage: no difference between the two (interaction between RC type (restrictive, non-restrictive) and modifier position (subject, object)

Conclusion

Information flow is another important factor in determining sentence complexity.

Background information is processed more easily early in a sentence. Foreground / new information is processed more easily later in a sentence.

On-Line Language Processor

CENTRAL PROCESSES

Semantics/Discourse
Syntax
Lexicon

Auditory	Visual

INPUT SIGNALS

Figurative Language

- Non-literal language

- Pragmatics: The interpretation of non-literal language

- Non-literal communication is ubiquitous
  Your mouth is sewer.

  Sarah: *I'm leaving you*
  John: *Who is he?*

Common Types of Figurative Language

(1) Metaphors

    ARGUMENTS ARE WARS:

    I demolished his argument.

    Your claims are indefensible.

    TIME IS MONEY:

    The flat tire cost me an hour.

    You're wasting time.

    This gadget will save you hours.

(2) Irony: When thoughts and deeds conflict

➢ Sarcasm:

Son: It's not that I don't want to work, it's just …

Mother: What kind of work did you have in mind? Watching Television? Listening to records? Playing the guitar? Driving the car?

(3) Metonymy: An easily perceived aspect of something represents the thing as a whole.

➢ Place for an institution

Washington has started negotiating with Moscow.

Wall Street is in a panic.

➢ Object for a user

The buses are on strike.

The sax has the flu today.

➢ Controller for the Controlled

Bush bombed Iraq.

Ozawa gave a terrible concert.

(4) Idioms: colloquial expressions:

➢ "spill the beans", "lose your marbles", "button your lip"

➢ Frozen complex lexical expressions?

John laid down the law.
The law was laid down by John.

John kicked the bucket.
*The bucket was kicked by John.

The shit hit the fan.
*The fan was hit by the shit.

(5) Indirect speech acts:

Requesting that someone shut the door

- Questioning the ability of someone to shut the door

  Can you shut the door?

- Questioning the willingness of the listener to perform the action

  Will you shut the door?

- Expressing the speaker's desire

  I would like the door shut.

- Questioning whether the act of shutting the door would impose on the listener

  Would you mind shutting the door?

- Making a statement about some relevant fact in the world

  It's cold in here.

Early Approaches to Interpreting Figurative Language

- Perhaps there is something syntactically odd to identify them?

  Bill tripped the light fantastic. (old idiomatic expression for "to dance")

- Perhaps figurative expressions include semantic violations

  - E.g., animate things cannot be subsets of inanimate things

  Your mouth is a sewer.

- The literal falsehood of a sentence might signal that it is a figurative expression.

  Your mouth is a sewer.

- **But, examples like:** *My husband is an animal.*

- The form and literal meaning of the linguistic expression by themselves offer no clues to figurativity.

Using **Coversational Cooperativity** to Signal Figurativity

My husband is an animal

  ➢ The literal meaning is odd: Why would the speaker say this? The literal meaning is not informative.

  ➢ The listener infers a non-literal meaning (her husband is crude, acts poorly towards her, etc.)

- Traditional view of processing figurative language:

  ➢ Compute the literal meaning first, and then if it is odd in some way (e.g., false, implausible, or violates one of **Grice's maxims**), then compute a figurative meaning.

Early View: Literal Interpretations are computed first

Bottom-up / Modularity hypothesis: Perhaps compute the literal meaning first, to see if that fits in the context. If not, an **implicature** is computed.

Therefore, it will take more time to compute non-literal language.

(The current evidence shows that this is unlikely to be the case: People use top-down information from the context to process all aspects of language.)

Early View: Literal Interpretations are computed first

Bottom-up / Modularity hypothesis: Perhaps compute the literal meaning first, to see if that fits in the context. If not, an **implicature** is computed.

Joe: Are you going to the big dance tonight?
Sue: Didn't you hear that Billy Smith will be there?

- Sue's answer means different things in different circumstances
  - ➢ If Sue likes Billy, she will likely want to go.
  - ➢ If Sue hates Billy, she will likely not want to go.

- The intended message behind Sue's statement is an **implicature**: the non-literal information that it conveys

Grice's Maxims: Informally deriving non-literal interpretations for sentences

- All speakers adhere to the cooperative principle:

  " Make your conversational contribution such as it is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged." (Grice, 1975, p.45)

- Cooperativity allows much to be left unsaid.

Grice's Maxims

Four conversational maxims for a cooperative speaker:

(1) Maxim of Quantity:

- Make your contribution as informative as is required
- Do not make your contribution more informative than is required

She sang the "Star Spangled Banner."

??She used her vocal tract to produce a series of sounds corresponding to the score of the "Star Spangled Banner."

(2) Maxim of Quality:

- Do not say that which you believe to be false
- Do not say that for which you lack evidence

Suppose Paul has just divulged a personal secret

??Paul is a fine friend.

Grice's Maxims

(3) Maxim of Relation:

- Say only what is relevant for the current purposes of the conversation.

At a genteel tea party

  A: *Mrs. Bigglesworth is an old bag.*

??B: *The weather has been quite delightful this summer.*

(4) Maxim of Manner:

- Be brief but avoid ambiguity or obscurity of expression.

  Joe: Are you going to the big dance tonight?

??Sue: Didn't you hear Billy Smith will be there?

Grice's Maxims

- As long as the speaker adheres to the cooperative principle, he/she can disobey the maxims intentionally.
  - ➢ Deliberate violation of a maxim can give rise to an **implicature**.
  - ➢ **Implicature**: exploiting the cooperative principle to convey more information than is actually contained in an utterance.

  - ➢ Hyperbole, sarcasm, understatement are all violations of Quality maxim.

Flouting Grice's Maxims

- Letter of recommendation for graduate school
  - ➢ Dear Sirs, Mr. X's command of English is excellent, his attendance at tutorials has been regular, and his wife is charming. – Yours, Professor Y.

    Violation of the maxim of quantity.

  - ➢ A: John doesn't seem to have a girlfriend these days.
    B: He's been driving up to New York every weekend.

    Violation of the maxim of relation and / or manner.

Using Grice's Maxims to Signal Figurativity

My husband is an animal

> ➢ The literal meaning violates the maxims of quantity (informativeness) and relation.

> ➢ The listener infers a non-literal meaning (her husband is crude, acts poorly towards her, etc.)

- Traditional view of processing figurative language:

> ➢ Compute the literal meaning first, and then if it is odd in some way (e.g., false, implausible, or violates one of Grice's maxims), then compute a figurative meaning.

But it Depends on the Context…

- Gibbs (1979)
  - ➤ Mrs. Smith was watering her garden one afternoon. She saw the housepainter was pushing a window open. She didn't understand why he needed to have it open. Worried that it might be difficult to get it open, she went over and politely asked, "**Do you have to open the window**?"
    - Paraphrase: Do you need to open the window?
  - ➤ One morning John felt too sick to go to school. The night before he and his friends got very drunk. Because of this he caught a bad cold. He was lying in bed when his mother stormed in. When she started to open the window, John groaned, "**Do you have to open the window**?"
    - Paraphrase: Do not open the window.

- Gibbs' results: No differences in RTs to a sentence whose literal meaning is appropriate vs. the same sentence when its implicated meaning is appropriate.

Local summary

- Much of communication is implicit (non-literal)
- We use the norms of cooperative conversation to help establish what is being conveyed implicitly.
- In processing figurative language
  - ➢ Sometimes we process the literal meaning first
    - If the context does not support a figurative interpretation
  - ➢ Other times we can directly access the figurative interpretation.
    - If the context allows us to activate enough semantic information to access the non-literal meaning directly
    - Top-down information from the context may make a non-literal interpretation the first that is available

Processing non-literal language

- Gricean Maxims: Cooperative conversation.
  - ➤ Violating a maxim leads to an implicature.
  - ➤ The Maxims are stated informally: It's hard to know how to apply them without a formal definition

- One theory of processing some non-literal language
  - ➤ **The Referential Theory**
    - Minimize discourse structure, including counting all Gricean implicatures.
    - Relies on violations of Gricean informativity in order to compute implicatures.

Referential theory (Crain & Steedman, 1985; Altmann & Steedman, 1988)

The referential theory was developed to account for the observation that the null context is not necessarily a neutral context: The null context might favor one interpretation over another.

E.g. MV/RR ambiguity:

(1) # The horse raced past the barn fell. (Bever, 1970)

The standard view in the literature was that the garden-path effect in (1) was due to a syntactic preference for the MV over the RR structure.

The referential theory

- **I: The principle of parsimony** (Crain & Steedman, 1985): A reading which carries fewer unsupported implicatures will be favored over one that carries more.

- **II: The principle of referential support** (Altmann & Steedman, 1988): An NP analysis which is referentially supported will be favored over one that is not.

  Referential theory explanation of the garden-path effect in (1): There are fewer unsupported *implicatures* (C&S refer to them as presuppositions, but this isn't quite the right term) in the MV structure than in the RR structure. This follows from the principle of parsimony.

Discourse structures in the mental model for the MV structure of "the horse raced":

1. A horse h_i;

Discourse structures in the mental model for the RR structure of "the horse raced":

1. A horse h_i;
2. A set of horses H of which h_i is a member;
3. One of this set, h_i, was raced somewhere;
4. None of the other members of the set H have the property in (3), that they were raced in the same way that h_i was raced.

There are three additional implicated structures in the RR structure, so the MV reading is preferred in the null context.

More general case:

When the definite article "the" is used with a head noun and a modifier (either before or after the noun), this implicates the existence of a set of nouns of which only one has the property indicated by the modifier.

This hypothesis relies on an implicature from a violation of a Grice's conversational maxim of quantity:

Maxim of quantity: Speakers should say as much as needed to be informative, without saying more than is necessary. (Grice, 1975)

Otherwise the speaker could simply have said "the horse fell" instead of "the horse raced past the barn fell."

Experimental tests of the principle of parsimony

1. **Crain & Steedman (1985): Bare plurals vs. definite plurals in the MV/RR ambiguity**

    Definite determiner ("the") + modifier: implicates the existence of a contrast set.

    Indefinites ("a", or bare plural) + modifier: weaker implication of the existence of a contrast set.

    E.g.,

    "The teacher that was taught by the Berlitz method"
    This implicates a set of teachers, one of whom was taught by the Berlitz method, and the others were not.

    "A teacher that was taught by the Berlitz method"
    This introduces a teacher, but has a weaker implication of for a set of teachers.

    "Teacher that were taught by the Berlitz method"
    This implicates a set of teachers that were taught by the Berlitz method. The contrast set (the teachers that were not taught by the Berlitz method) is more weakly implicated.

Experimental tests of the principle of parsimony

1. **Crain & Steedman (1985): Bare plurals vs. definite plurals in the MV/RR ambiguity**

    Speeded grammaticality task:

    (2) a. Definite plural, Plausible subject
    The teachers taught by the Berlitz method passed the test.
    b. Bare plural, Plausible subject
    Teachers taught by the Berlitz method pass the test.
    c. Definite plural, Implausible subject
    The children taught by the Berlitz method passed the test.
    d. Bare plural, Implausible subject
    Children taught by the Berlitz method pass the test.

Crain & Steedman (1985): Bare plurals vs. definite plurals in the MV/RR ambiguity

Results:

1. Bare plurals were accepted as grammatical more than definite plurals;

2. Implausible subject sentences were accepted as grammatical more than plausible subject sentences.

Ni, Crain & Shankweiler (1996): "only" vs. "the" in MV/RR ambiguity

"the" + N + modifier: implication of a contrast set of N's.
All modifiers of a head Noun are generally taken to be contrastive following the determiner "the".

The focus operator **"only"** *always needs* a contrast set:

(5) In New Haven only Willoughby's coffee is really good.

(5) is true if: (a) the coffee at Willoughby's is good; and (b) the coffee everywhere else in New Haven is not good.

This is an implicit comparison: a contrast set. If no contrast set is explicitly mentioned in the discourse, then one has to be constructed.

Ni, Crain & Shankweiler (1996): "only" vs. "the" in MV/RR ambiguity

Experiments 1 and 2: MV/RR in self-paced reading and eye-tracking

(6)
a. The, ambiguous (adjective)
The (wealthy) businessmen loaned money at low interest were told to record their expenses.
b. Only, ambiguous
Only (wealthy) businessmen loaned money at low interest were told to record their expenses.

(7) Unambiguous controls:
The/Only (new) vans stolen from the parking lot were found in a back alley.

Ni, Crain & Shankweiler (1996): "only" vs. "the" in MV/RR ambiguity

Predictions:

Minimal attachment: reanalysis effect at "were told…" in all four versions.

Referential theory: reanalysis effect at "were told…" in (a), (c) and (d) versions, but not in the (b) version.

In (a) and (c), this prediction is because of the use of the definite determiner "the" and a potential modifier of the head noun (the RR reading): more implicated discourse structure.

Ni, Crain & Shankweiler (1996): "only" vs. "the" in MV/RR ambiguity

(b) Only businessmen loaned money at low interest were told to record their expenses.

In (b), the focus operator "only" needs to find a contrast set for its head noun "businessmen":

Two options:

1. The MV interpretation: Create a contrast set out of thin air: businessmen as opposed to other types of men / people

2. The RR interpretation: Break the set of businessmen into two subsets. The RR modification of businessmen is one way of instantiating this option: Contrast set already given.

Therefore, the referential theory (the principle of parsimony) predicts that people will follow the RR reading.

Ni, Crain & Shankweiler (1996): "only" vs. "the" in MV/RR ambiguity

(d) Only wealthy businessmen loaned money at low interest were told to record their expenses.

In (d), the contrast set for "only" is provided by the adjective "wealthy". There is no need for another contrast set at "loaned", so the MV structure is predicted to be preferred.

Ni, Crain & Shankweiler (1996): "only" vs. "the" in MV/RR ambiguity

Summary: condition (b) is predicted to pattern like the unambiguous controls. The others are predicted to show reanalysis effects when compared to the unambiguous controls.

Ni, Crain & Shankweiler (1996): "only" vs. "the" in MV/RR ambiguity

Self-paced reading results:

Graph removed for copyright reasons.

The predictions of the referential theory were ratified.

Ni, Crain & Shankweiler (1996): "only" vs. "the" in MV/RR ambiguity

Eye-tracking results:

Graph removed for copyright reasons.

The predictions of the referential theory were ratified.

Processing Semantic and Pragmatic information: Conclusions day 1

1. Information flow in sentence processing
   - Grammatical constructions have **functions.** E.g., Restrictive modifier: pick an item from a set.
   - Background information is processed more quickly early in a sentence.

2. Much of communication is implicit (non-literal)
   - We use the norms of cooperative conversation to help establish what is being conveyed implicitly.
   - In processing figurative language, sometimes we process the literal meaning first; Other times we can directly access the figurative interpretation.

3. One theory of processing some non-literal language
   - **The Referential Theory**
     - Minimize discourse structure, including counting all Gricean implicatures.
     - Relies on violations of Gricean informativity in order to compute implicatures.