

Language comprehension

Lecture 6: Pragmatics and context in sentence comprehension

9.591; 24.945

October 25, 2004

Ted Gibson

9.591 Course so far

- Lecture 1: Experimental methods; Informational constraints affecting sentence comprehension: Lexical frequency, plausibility, context and syntax; Modularity in sentence comprehension.
- Lecture 2: Resources and sentence complexity. The complexity of unambiguous sentences.
- Lecture 3: Working memory and sentence comprehension.
- Lecture 4: Ambiguity resolution: Resources; structural frequencies.
- Lecture 5: Experience / frequency and ambiguity resolution; the serial / parallel issue.
- **Lecture 6: Pragmatics and context in sentence comprehension**

Sentence processing: Recap thus far

- Multiple factors are involved in processing unambiguous sentences and in ambiguity resolution.
 - **Discourse context**
- What is the time course of information integration?
 - Modular (syntax-first)? Or non-modular?
- Is the parser serial or constrained parallel?

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 presuppositions 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 presuppositions (this is C&S's term: a better term is *implicatures*) in the MV structure than in the RR structure. This follows from the principle of parsimony.

To do today:

- Evaluating the referential theory (Crain & Steedman, 1985; Altmann & Steedman, 1988):
 - Tests of the principle of parsimony in null contexts
 - Tests of the principle of referential support in contexts
 - Generalizing referential theory to the processing of unambiguous sentences (Grodner, Gibson & Watson, in press)
- Pragmatic processing: Implicatures and on-line sentence processing (Sedivy et al.)
 - When do people compute contrast sets?

Presuppositions in the mental model for the MV structure of “the horse raced”:

1. A horse h_i ;

Presuppositions 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 presuppositions 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**

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 (no means given: only stats):

1. Bare plurals were accepted as grammatical more than definite plurals;
2. Implausible subject sentences were accepted as grammatical more than plausible subject sentences.

Spivey-Knowlton & Sedivy (1995): V-NP-PP

- a. VP attachment, def / indef NP
The fireman | smashed down | **the / a door** | with **the / a heavy axe** | but smoke | ...
- b. NP attachment, def / indef NP
The fireman | smashed down | **the / a door** | with **the / a rusty lock** | but smoke | ...

Indefinite singular NPs (“a door”) vs. definite singular NPs (“the door”) in self-paced reading, region by region.

All items involve the preposition “with”

Spivey-Knowlton & Sedivy (1995): V-NP-PP

Minimal Attachment prediction:

VP attachment preference, for both definite and indefinite NPs.

Referential theory prediction:

An interaction: VP attachment in definite conditions; but no difference in indefinite conditions.

Spivey-Knowlton & Sedivy (1995): V-NP-PP

Experiment 1 Results:

RTs in msec during the disambiguating PP ("with a/the heavy
axe/rusty lock")

	Def NP	Indef NP
VP attach	728	677
NP attach	835	724

RTs in msec in the region following the disambiguating PP ("but
smoke")

	Def NP	Indef NP
VP attach	661	638
NP attach	704	698

Spivey-Knowlton & Sedivy (1995): V-NP-PP

Experiment 1 Results:

1. VP attachment was faster than NP attachment in the PP region. This fits Minimal Attachment (and argument preference more generally. Cf. Schutze & Gibson, 1999)
2. The indefinite conditions were faster than the definite conditions. This may simply reflect string length differences: "a" is shorter than "the".

Spivey-Knowlton & Sedivy (1995): V-NP-PP

Experiment 1 Results:

3. There was a suggestion of an interaction between definiteness and NP/VP attachment in the region following the disambiguating region, but this was not significant in the items analysis.

Overall: The results seem to support Minimal Attachment over the referential theory.

Spivey-Knowlton & Sedivy (1995): V-NP-PP

Alternative possibility: Perhaps the verbs that were used in Experiment 1 are lexically biased to prefer VP attachment of “with” PPs. Other verbs might not be biased in this way.

Verbs in Experiment 1: eventive, action verbs.

New set of verbs in Experiment 2: stative and perception verbs

Spivey-Knowlton & Sedivy (1995): V-NP-PP

Experiment 5 materials:

a) VP attachment, def / indef NP

The salesman | glanced at | **the / a customer** |
with **suspicion** | and then | ...

b) NP attachment, def / indef NP

The salesman | glanced at | **the / a customer** |
with **ripped jeans** | and then | ...

Spivey-Knowlton & Sedivy (1995): V-NP-PP

Experiment 5 Results:

RTs in msec during the disambiguating PP (“with suspicion / ripped jeans”)

	Def NP	Indef NP
VP attach	707	744
NP attach	770	676

No overall NP or VP attachment preference, but a reliable interaction between attachment site and definiteness, as predicted by referential theory.

Spivey-Knowlton & Sedivy (1995): V-NP-PP

Conclusions:

Multiple constraints affect people's initial preferences in resolving syntactic ambiguity.

1. Verb type / argumenthood status of the PP / lexical argument structure frequencies
2. Definiteness of the NPs

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

The focus operator “only” 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 (adjective)
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).

This prediction is because of the use of the definite determiner “the” and a potential modifier of the head noun (the RR reading).

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

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

In this sentence, the focus operator “only” needs to find a contrast set for its head noun “businessmen”:

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

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

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

In this structure, 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: “only” with no adjective 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:

Two graphs 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.

Ni, Crain & Shankweiler (1996): Experiments 3 and 4: "only" vs. "the" in PP attachment

(7) a. VP attachments

The men painted the / only doors with new brushes before the festival.

b. NP attachments

The men painted the / only doors with large cracks before the festival.

Predictions:

MA: VP attachment preference

Referential theory:

VP attachment preferred for definite determiner "the".

NP attachment preferred for focus operator "only"

Ni, Crain & Shankweiler (1996): Experiments 3 and 4: "only" vs. "the" in PP attachment

Graph removed for copyright reasons.

Tests of the principle of referential support

Suppose that there is no difference in the number of unsupported presuppositions for either of two readings.

E.g., Consider (1) in the context of a horse, and a set of horses, only one of which was raced past the barn.

(1) The horse raced past the barn fell.

The principle of parsimony does not apply: no unsupported presuppositions in either interpretation.

Tests of the principle of referential support

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

An NP is referentially supported if we can pick out the appropriate object of the referring expression from the context.

Avoid infelicitous reference!

In the above context, the RR analysis is referentially supported, but the MV is not. The MV structure includes a reference to "the horse", but does not specify which horse, so it is not referentially supported.

Tests of the principle of referential support

1. Crain & Steedman (1985): V-NP-CP ambiguity

Clausal argument of verb (complement clause / CC) or modifier of NP (relative clause/ RC)

a. Contexts: One or two NP referents

A psychologist was counseling {(two women) | (a man and a woman)}. He was worried about one of them but not the other.

b. Target sentence: RC/CC

The psychologist told the woman that he was having trouble with {(to leave) / (her husband)}.

Tests of the principle of referential support

Predictions:

Minimal Attachment: VP attachment preference independent of context.

Referential theory: CC preference in one referent context; RC preference in two referent context

Task: speeded grammaticality.

Results: The interaction predicted by referential support was observed:

	2 referents	1 referent
VP attachment	46%	78%
NP attachment	88%	50%

Some experiments that failed to find that contexts helped

Ferreira & Clifton (1986): V-NP-PP ambiguity

Neutral or NP supportive contexts:

- a. VP-attachment: Sam loaded the boxes on the cart before his coffee break.
- b. NP-attachment: Sam loaded the boxes on the cart onto the van.

No effect of context: The VP disambiguations were read faster independent of context.

Some experiments that failed to find that contexts helped

Ferreira & Clifton (1986): V-NP-PP ambiguity

But problem: The verbs in the study (mostly “load” and “place”) may be too biased to VP attachments to see any context effects (Spivey-Knowlton & Tanenhaus, 1994).

Some experiments that failed to find that contexts helped

Ferreira & Clifton (1986): MV/RR ambiguity

Sentences in isolation or in RR supporting contexts:

- a. MV: The editor played the tape and agreed the story was big.
- b. RR: The editor played the tape agreed the story was big.

Contexts: One editor or two editors.

Results: No effect of context: The MV disambiguations were read faster independent of context.

Some experiments that failed to find that contexts helped

Ferreira & Clifton (1986): MV/RR ambiguity

But problem: The contexts weren't biased enough towards the RR structure (Spivey-Knowlton & Tanenhaus, 1994).

In particular, there were additional unsupported presuppositions in the RR structures that were not made explicit in the contexts.

The context did not provide the information that one of the editors was played a tape, and the other was not.

Some experiments that failed to find that contexts helped

3. Murray & Liversedge (1994): MV/RR ambiguity, eye-tracking

a. Context: One vs. two referents

Two people were discussing the food that had been prepared at the barbecue. One was a guest that enjoyed meat and the other was a caterer / guest who was a vegetarian.

b. Target

MV: The guest grilled the steak and said it tasted nice.

RR: The guest (who was) grilled the steak said it tasted nice.

Some experiments that failed to find that contexts helped

Results:

The disambiguating region was read slower for the RR compared to either the MV or the unambiguous RR, regardless of context.

But problem:

The contexts aren't biased enough toward the RR structures. The context does not specify that someone had grilled a steak for exactly one of the guests. According to referential theory, this discourse structure needs to be added when the RR structure is processed.

Spivey-Knowlton & Tanenhaus (1994)

Experiment 1:

One referent context:

An actress and the producer's niece were auditioning for a play. The director selected the actress but not the niece.

Two referents context:

Two actresses were auditioning for a play. The director selected one of the actresses but not the other.

Target:

The actress (who was) selected by the director believed that her performance was perfect.

Spivey-Knowlton & Tanenhaus (1994)

Results:

Graph removed for copyright reasons.

Spivey-Knowlton & Tanenhaus (1994)

Results:

Significant interaction between context and target sentences in the disambiguating region “by the director”:

Ambiguous targets were read more slowly than unambiguous controls in the one referent condition, but no differences in the two referent condition.

This is as predicted by referential support.

Monitoring visual eye-movements while listening to spoken instructions

“Put the frog on the napkin into the box.”

Photo removed for copyright reasons.

Monitoring visual eye-movements while listening to spoken instructions

“Put the frog on the napkin into the box.”

Two frog context: No looks to the incorrect target (the second napkin)

Photo removed for copyright reasons.

One frog context: Many looks to the incorrect target (the second napkin)

Trueswell et al. 1999: Children age 5 can't do the task

Four graphs removed for copyright reasons.

Trueswell et al. 1999

- Trueswell et al.'s conclusion:
 - Lexical frequency information dominates contextual information

But why?

Wexler: Alternative explanation: Kids at this age don't know the meaning of the word "the".

Fits with developmental data.

Generalizing referential theory to the processing of unambiguous sentences

(Grodner, Gibson & Watson, in press)

The referential theory (parsimony, referential support) as stated applies only in the resolution of ambiguity.

Three ways to extend Referential Theory

(1) Ambiguity Only Hypothesis

The discourse is consulted only in the face of ambiguity. The processing mechanism interprets an ambiguous utterance so as to make the background assumptions of the utterance consonant with a relevant model of the discourse context.

Generalizing referential theory to the processing of unambiguous sentences

(2) Weakly-Interactive Mental Models Hypothesis (Modular: syntax first):

Sentences are parsed using intrasentential criteria, such as syntactic knowledge. The resultant analysis (or analyses in the case of ambiguity) is then evaluated against the context, and changes are incrementally made to the current discourse model. These changes can incur costs that interfere with interpretive processes and lead to comprehension difficulty.

Generalizing referential theory to the processing of unambiguous sentences

(3) Strongly-Interactive Mental Models Hypothesis (Non-modular):

The discourse model is constantly updated and accessed in the comprehension of a sentence. Sometimes the sentence causes the construction of discourse structure. Other times the discourse model directs interpretive processes and projects syntactic structures.

Materials to test the hypotheses: Restrictive and non-restrictive RCs

- Restrictive RCs: usually function to identify a set. In a null context, they often implicate a contrast set.
 - Non-restrictive RCs: additional information about the head. No implication of a contrast set.
- (4) a. Restrictive RC: The boy that studied for the exam aced the test.
- b. Non-restrictive RC: Mary, who studied for the exam, aced the test.

Restrictive and non-restrictive RCs

Non-restrictive RCs can modify nominal heads that do not permit contrast:

(5)

- a. My father, who ate ham this morning, became extremely ill.
- b. The sun, which rises in the east, can be used to orient oneself.

Restrictive RCs can't:

(6)

- a. *My father that ate ham this morning became extremely ill.
- b. *The sun that rises in the east can be used to orient oneself.

Restrictive and non-restrictive RCs

Observation: Restrictive RCs involve more discourse structure than non-restrictive RCs: a contrast set. Therefore, if referential theory applies in unambiguous structures, a restrictive RC should be harder to process in a null context, other factors being equal.

Restrictive and non-restrictive RCs

(7) Null Context

a. Restrictive RC

A postman that a dog bit on the leg needed seventeen stitches and had a permanent scar from the injury.

b. Non-Restrictive RC

A postman, who a dog bit on the leg, needed seventeen stitches and had a permanent scar from the injury.

Restrictive and non-restrictive RCs

Predictions:

Ambiguity only: no differences.

Both strong and weak mental models hypotheses predict that the restrictive RC should be slower to process than the non-restrictive RC.

Prediction of strong mental models hypothesis

In certain specific situations, the discourse context can cause people to expect certain syntactic structures.

One instance where this may be the case: processing a definite NP (e.g., “the postman”) when there are two potential referents for the head noun in the discourse (e.g., two postmen).

In this situation, we may **expect** modification.

Contextually supported RCs

(8) Supportive Context

a. Restrictive RC

A vicious guard dog bit a postman on the leg and another postman on the arm.

The postman that the dog bit on the leg needed seventeen stitches and had a permanent scar from the injury.

b. Non-Restrictive RC

A vicious guard dog bit a postman and a garbage man.

The postman, who the dog bit on the leg, needed seventeen stitches and had a permanent scar from the injury.

Contextually supported RCs

Prediction of strong mental models hypothesis:
restrictive RCs should be **faster** than non-restrictive
RCs in a supportive context.

The ambiguity-only hypothesis and the weakly
interactive mental model hypothesis do not make this
prediction.

Experiment 1 materials

Null Context, Restrictive RC

A postman **that a dog bit on the leg** needed seventeen stitches...

Null Context, Non-Restrictive RC

A postman, **who a dog bit on the leg**, needed seventeen stitches...

Supportive Context, Restrictive RC

A vicious guard dog bit a postman on the leg and another postman on the arm.

The postman **that the dog bit on the leg** needed seventeen stitches...

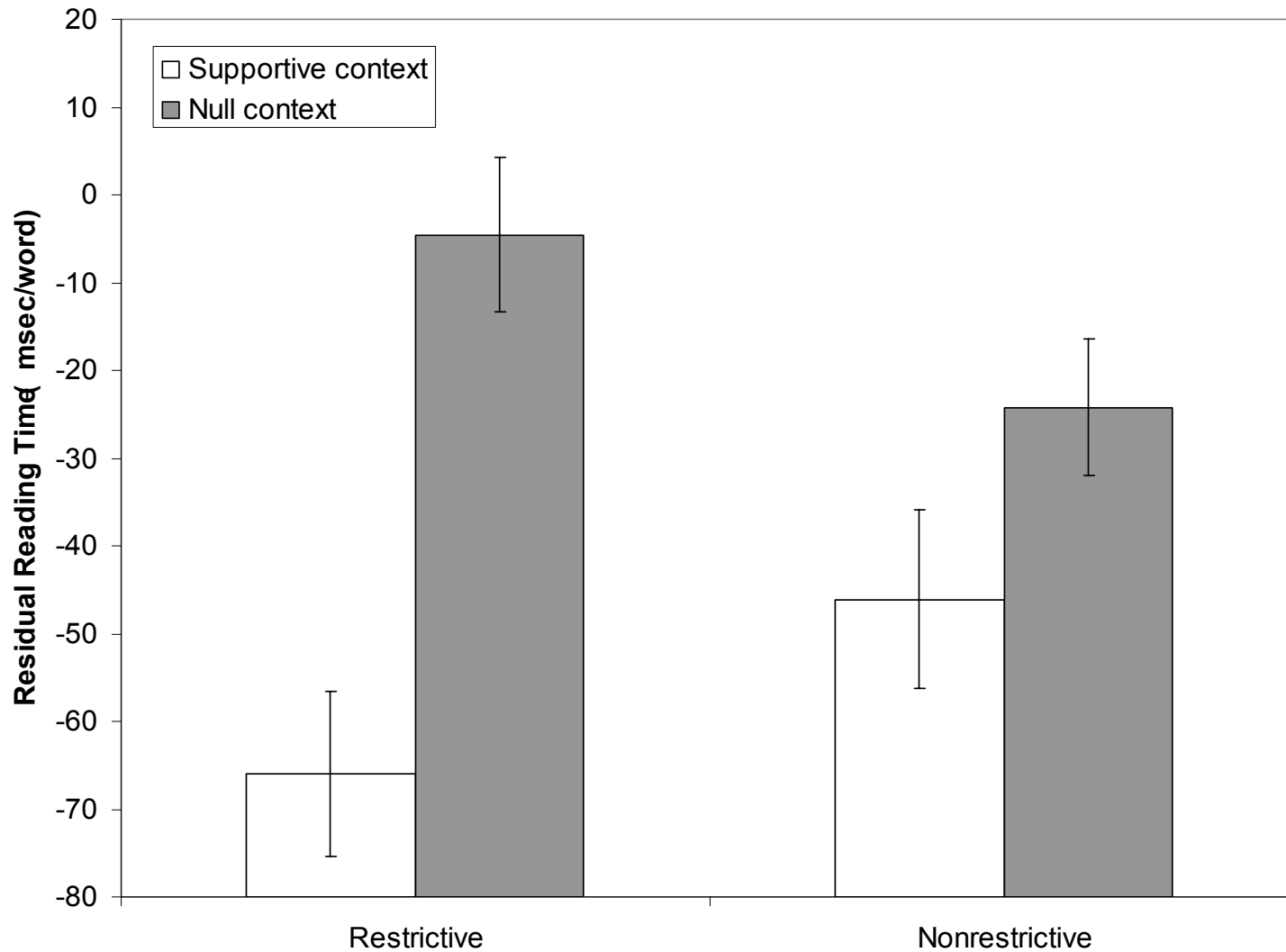
Supportive Context, Non-Restrictive RC

A vicious guard dog bit a postman and a garbage man.

The postman, **who the dog bit on the leg**, needed seventeen stitches...

Results: Discourse complexity experiment

Reaction times during the noun and verb of the RCs ("dog bit")



Results: Discourse complexity experiment

- (1) In the null context, restrictive RCs were read slower than non-restrictive RCs.
- (2) In a supporting context, restrictive RCs were read faster than non-restrictive RCs

These results were as predicted by the strongly-interactive mental models hypothesis, but not by the other hypotheses.

Conclusions of discourse complexity experiment (Grodner, Gibson & Watson, in press)

- (1) A referential theory / mental models theory applies in unambiguous structures, in addition to ambiguous structures.
- (2) The discourse context may sometimes be strong enough to guide syntactic structure building.

Pragmatic processing: Implicatures and on-line sentence processing

- When do people compute contrast sets?
- Some determiners seem to need contrast sets more than others: “only” vs. “the” vs. “a”

Contrastive Inferences

Q: *What time is it?*

A: *Some people are already leaving.*

→ It's late.

Q: *How is the party?*

A: *Some people are already leaving.*

→ The party isn't very good.

- When are inferences computed?
- What aspects of the context enter into their computation?

Dependency Between Restrictive Modification and Contextual Contrast

Can you pass Tim the tall cup?

$\exists!x[\text{cup}(x) \ \& \ \text{tall}(x)]$

reference set

$\rightarrow \exists x[\text{cup}(x) \ \& \ \neg \text{tall}(x)]$

contrast set

Sedivy, Chambers, Tanenhaus, & Carlson (1999)

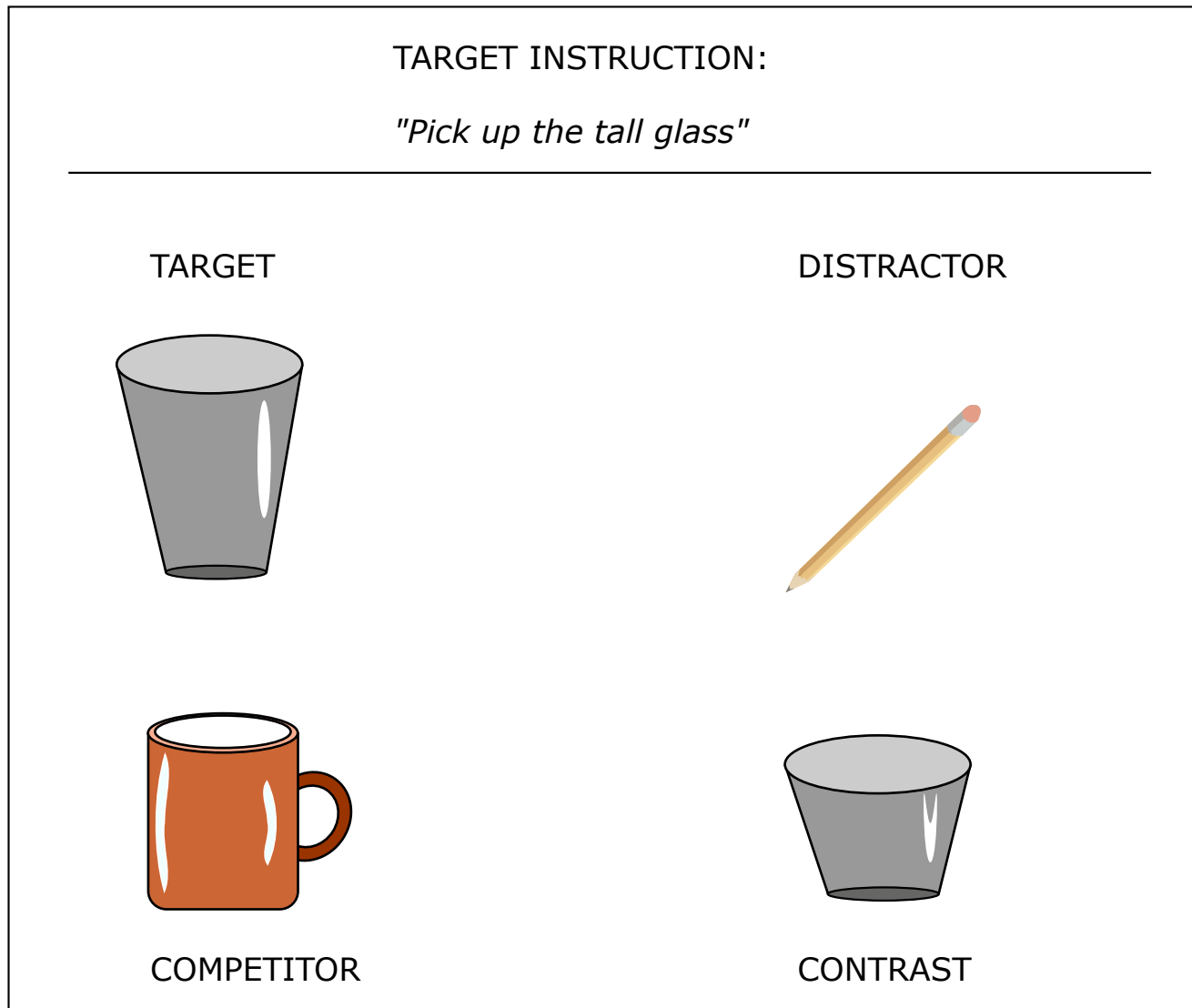


Figure by MIT OCW.

•Contrast Effect: Eye-movements converge more quickly on the target and there are fewer looks to the competitor in the presence of a contrast set.

Two Classes of Explanation for Contrastive Inferences

(1) Form-Based Account

- Contrastive inference is closely tied to conventional meaning of restrictively modified NPs or to the lexical class of the modifier.
 - Scalar adjectives contain a variable assigned by a contextually relevant comparison class (Seigel, 1980; Bierwisch, 1987)
- Minimizes the amount of information that is accessed in making contrastive inferences

Two Explanations for Contrastive Inferences

(2) Gricean Account

- Contrastive inferences arise because the use of a restrictive modifier is embedded in a collaborative communicative context.
- Quantity-2: Don't make your contribution more informative than is required for the purposes of the present exchange.
- The hearer notes that the speaker chose a modified form rather than an unmodified form to refer to an entity. The inclusion of the modifier is most easily made informative by attributing to it a distinguishing function.

Tests of the theories

- Testing the form-based account: The contrast effect should disappear if a non-scalar adjective is used. E.g., a color adjective, or a material adjective.
- Colors: “pick up the blue cup”, with a blue and red cup in the display. (as well as a competitor blue object, in order to control for the fact that people are incremental in their eye-gazes.)
- Results from colors:
The contrast effect disappears!
- Is this support for the form-based account?
 - Yes, but there is an alternative Gricean account.

Tests of the theories

- Surprising result for the form-based account:
- Materials: “pick up the plastic cup”, with a plastic cup and glass cup in the display.
- Results from materials:
The contrast effect re-appears!
- This contradicts the prediction of the form-based account

New Gricean theory (Sedivy, 2003; cf. Levinson, 2000)

- Quantity-2: Don't make your contribution more informative than is required for the purposes of the present exchange.
- The hearer notes that the speaker chose a modified form rather than the **simple, default form** to refer to an entity. The inclusion of the modifier is most easily made informative by attributing to it a distinguishing function.

(1) **Neo-Gricean View (Conservative)**

Early contrastive-inferences are only sensitive to whether or not the speaker elaborates on a default form. (cf. Levinson, 2000)

- The baseline is the default form: the way that people would describe the situation with no contrasting information.

New Gricean theory (Levinson, 2000)

- Differences in default forms:
 - Colors are often produced along with the head noun in describing an object (Sedivy, 2003).
 - Materials and scalar adjectives are not.
- Thus, the presence of a material or scalar adjective provides suggestive information to the listener that there is a contrasting object in the relevant dimension. Colors do not provide this information.

Predictions of the neo-Gricean view

1. If a color term is not normally produced when describing an object, then the contrast effect should re-emerge.

Sedivy (2003): “Pick up the yellow banana”, in the context of a yellow banana and a green banana

Predictions of the neo-Gricean view

2. If the listener knows that the speaker is not reliable in his / her productions, then looks to the contrasting elements may disappear.

Grodner et al. (2003): This prediction is realized.

The Dangers of a Gricean World

Types of information potentially admitted in determining an expected referential form.

- Intrinsic properties of the target referent
 - Properties of other referents in the discourse context
 - The reliability of a speaker
 - The intentions of a speaker
 - Shared background assumptions
 - Expectations of goals of a communicative situation
 - Expectations about alternative forms
 - And so on ...
-
- Flexible but potentially slow. An expeditious processing system might only attend to some of these in the earliest stages of interpretive processing.