#### EMPIRICAL INVESTIGATIONS OF ANAPHORA

Massimo Poesio Università di Trento and University of Essex

Vilem Mathesius Lectures Praha, 2007

#### **Plan of the series**

- Yesterday: Annotating context dependence, and particularly anaphora
- Today: Using anaphorically annotated corpora to investigate local & global salience ('topics')
- Tomorrow: Using anaphorically annotated corpora to investigate anaphora resolution

#### **Today's lecture**

- Quick intro to the Grosz / Sidner framework for studying salience in discourse
- An investigation of local salience
- Local salience and demonstratives
- (If time allows): global salience

## **Topic / focus: the Prague view**

- Kind of foolish to come here to talk about topic / focus
- All I want to say is that in these lectures while I use the term 'topics' in more or less the sense that it is used in Prague, the term 'focus' is used as equivalent to topic ...

## Topic / `Focusing' effects in Linguistics and Psychology

- In Linguistics, the notion of `topic' used in explanations of
  - `topic marking' in languages like Japanese
  - zero pronominalization
- In Psychology, the notion of `focus' used to account for ambiguity reduction effects and shorter reading times
  - The little puppy trod on the wasp.
  - The puppy was very upset.
  - ?? It started to buzz furiously.

# The Grosz / Sidner framework for studying salience and coherence

- According to Grosz and Sidner, discourses have two levels of organization:
  - At a GLOBAL level, discourses are divided into SEGMENTS that must be coherent at the INTENTIONAL level (= intentions must be clearly related)
  - At a LOCAL level, segments must be coherent at the ENTITY level (= successive utterances must be about the same entities)

# The Grosz & Sidner theory of discourse structure

- As we will see, there are plenty of aspects of the theory that need to be clarified
- We will begin by looking at the claim at the LOCAL level
- Then if time permit we will look at the GLOBAL level
- Dissenting views
  - Knott et al: for certain types of texts, entity coherence is the determining factor at the global level

# Centering Theory (Grosz, Joshi and Weinstein, 1995)

- The component of Grosz and Sidner's theory of attention and coherence in discourse concerned with LOCAL effects (within discourse segments)
- Its aim is to make cross-linguistically valid claims about which discourses are easier to process, abstracting away from specific algorithms

### Claims of the theory, I: Entity coherence

- Discourses without a clear 'central entity' feel less coherent
  - (1) a. John went to his favorite music store to buy a piano.
    - b. He had frequented the store for many years.
    - c. He was excited that he could finally buy a piano.
    - d. He arrived just as the store was closing for the day.
  - (2) a. **John** went to his favorite music store to buy a piano.
    - b. It was a store John had frequented for many years.
    - c. He was excited that he could finally buy a piano.
    - d. It was closing just as John arrived.

Add examples of utterances and CFs!

## **Concepts and definitions, I**

- Every UTTERANCE U in a discourse (segment) DS updates the LOCAL FOCUS - a PARTIALLY RANKED set of discourse entities, or FORWARD-LOOKING CENTERS (CFs)
- An utterance U in discourse segment DS updates the existing CF set by replacing it with the set of CFs REALIZED in U, CF(U,DS) (usually simplified to CF(U))
- The most highly ranked CF realized in utterance U is CP(U)

(1) u1. Susan gave James a pet hamster.

CF(u1) = [Susan,James,pet hamster]. CP(u1) = Susan

(2) u2. She gave Peter a nice scarf.

CF(u2) = [Susan,Peter,nice scarf]. CP(u2) = Susan

#### **Concepts and Definitions,II:** The CB

 Constraint 3: The BACKWARD-LOOKING CENTER of utterance U<sub>i</sub>, CB(U<sub>i</sub>), is the highest-ranked element of CF(U<sub>I-1</sub>) that is realized in U<sub>i</sub>

#### **The CB: Examples**

(1) u1. Susan gave James a pet hamster.

CF(u1) = [Susan,James,pet hamster]. CB = undefined CP=Susan

(2) u2. She gave Peter a nice scarf.

CF(u2) = [Susan, Peter, nice scarf]. CB=Susan. CP=Susan

NB: The CB is not always the most ranked entity of the PREVIOUS utterance

(2') u2. He loves hamsters.

CF(u2) = [James]. CB=James. CP=James

- ... or the most highly ranked entity of the CURRENT one
- (2") u2. Peter gave her a nice scarf.

CF(u2) = [Peter,Susan, nice scarf]. CB=Susan. CP=Peter

#### **Constraint 1**

CONSTRAINT 1 (STRONG): All utterances of a segment except for the first have exactly one CB

CB UNIQUENESS: Utterances have at most one CB

ENTITY CONTINUITY: For all utterances of a segment except for the first,  $CF(U_i) \cap CF(U_{i-1}) \neq \emptyset$ 

CONSTRAINT 1 (WEAK): All utterances of a segment except for the first have AT MOST ONE CB

# Claims of the theory, II: Local salience and pronominalization

- Grosz et al (1995): the CB is also the most salient entity. Texts in which other entities (but not the CB) are pronominalized are less felicitous
  - (1) a. Something must be wrong with John.
    - b. He has been acting quite odd.
    - c. He called up Mike yesterday.
    - d. John wanted to meet him quite urgently.
  - (2) a. Something must be wrong with John.
  - b. He has been acting quite odd.
  - c. He called up Mike yesterday.
    - d. He wanted to meet him quite urgently.



#### Claims of the theory, III: Preserving the ranking

- Discourses without a clear 'central entity' feel less coherent
  - (1) a. John went to his favorite music store to buy a piano.
    - b. He had frequented the store for many years.
    - c. He was excited that he could finally buy a piano.
    - d. He arrived just as the store was closing for the day.
  - (2) a. **John** went to his favorite music store to buy a piano.
    - b. It was a store John had frequented for many years.
    - c. He was excited that he could finally buy a piano.
    - d. It was closing just as John arrived.

#### **Transitions**

 Grosz et al proposed that the load involved in processing an utterance depends on whether that utterance preserves the CB of the previous utterance or not, and on whether CB(U) is also CP(U). They introduce the following classification:

CENTER CONTINUATION:  $U_i$  is a continuation if  $CB(U_i) = CB(U_{i-1})$ , and  $CB(U_i) = CP(U_i)$ 

CENTER RETAIN:  $U_i$  is a retain if  $CB(U_i) = CB(U_{i-1})$ , but  $CB(U_i)$  is different from  $CP(U_i)$ 

CENTER SHIFT: U<sub>i</sub> is a shift if  $CB(U_i) \neq CB(U_{i-1})$ 

#### **Massimo Poesio:**

Note that you need to establish the CB first – see Walker et al 1994, Kameyama 1998, etc.

#### **Utterance classification**

(0) u0. Susan is a generous person.

CF(u0) = [Susan] CB = undefined CP = Susan.

(1) u1. She gave James a pet hamster.

CF(u1) = [Susan,James,pet hamster]. CB = Susan CP=Susan

#### CONTINUE:

(2) u2. She gave Peter a nice scarf.

CF(u2) = [Susan,Peter,nice scarf]. CB=Susan. CP=Susan CONTINUE

### **Utterance classification, II**

(0) u0. Susan is a generous person.

CF(u0) = [Susan] CB = undefined CP = Susan.

(1) u1. She gave James a pet hamster.

CF(u1) = [Susan,James,pet hamster]. CB = Susan CP=Susan

#### SHIFT:

(2') u2. He loves hamsters.CF(u2) = [James]. CB=James. CP=James SHIFT

#### **Utterance classification, III**

(0) u0. Susan is a generous person.

CF(u0) = [Susan] CB = undefined CP = Susan.

(1) u1. She gave James a pet hamster.

CF(u1) = [Susan,James,pet hamster]. CB = Susan CP=Susan

**RETAIN:** 

(2") u2. Peter gave her a nice scarf.

CF(u2) = [Peter,Susan, nice scarf]. CB=Susan. CP=Peter RETAIN

#### Rule 2

RULE 2: (Sequences of) continuations are preferred over (sequences of) retains, which are preferred over (sequences of) shifts.

#### **Summary of the claims**

CONSTRAINT 1: All utterances of a segment except for the first have exactly one CB

RULE 1: if any CF is pronominalized, the CB is.

RULE 2: (Sequences of) continuations are preferred over (sequences of) retains, which are preferred over (sequences of) shifts.

## A parametric theory

- Central notions like 'utterance', previous utterance', `ranking', and `realization', left underspecified: they are PARAMETERS
- A lot of work has gone into identifying possible ways of setting these parameters
- Lots of empirical work has tested particular claims of the theory, but always with respect to a particular setting; comparisons between two ways of setting a specific parameter also exist
- No previous study has attempted to consider systematically variations in the instantiations of multiple parameters

#### **Goals of this work**

- To consider systematically the effect of parameter variation on the theory's claims
  - (Possibly identifying other parameters)
- To verify the extent to which texts follow Centering preferences
  - (Possibly, to identify other factors that may play a role)

#### **Outline of the talk**

- Centering's Parameters
- Using a corpus to test Centering
- Vanilla instantiation
- Other instantiations
- Discussion

### The parameters of the theory

- Grosz et al do not provide algorithms for computing any of the notions used in the basic definitions:
  - UTTERANCE
  - PREVIOUS UTTERANCE
  - REALIZATION
  - RANKING
  - What counts as a 'PRONOUN' for the purposes of Rule 1? (Only personal pronouns? Or demonstrative pronouns as well? What about second person pronouns?)
- One of the reasons for the success of the theory is that it provides plenty of scope for theorizing ...

#### **Utterance and Previous Utterance**

- Originally, utterances implicitly identified with sentences. Kameyama (1998) and others suggested to identify utterances with finite clauses.
- If utterances identified with sentences, the previous utterance is generally easy to identify. But if utterances are identified with finite clauses, there are various ways of dealing with cases like:
  - (u1) John wanted to leave home (u2) before Bill came home.
    (u3) He would be drunk as usual.
  - KAMEYAMA: PREV(u3) = u2.
  - SURI and MCCOY: PREV(u3) = u1

### Realization

- A basic question is whether entities can be 'indirectly' realized in utterances by an associate (as in Sidner's algorithm)
  - (u1) John walked towards the house.
  - (u2) THE DOOR was open.
- A second question is whether first and second person entities are realized:
  - (u1) Before you buy this medicine,
  - (u2) you should contact your doctor.
- Realization greatly affects Constraint 1.

#### Massimo Poesio:

A problem for Strube: how do you account for all that evidence about subject assignment in English?

## Ranking

- GRAMMATICAL FUNCTION (Kameyama 1986, Grosz Joshi and Weinstein 1986, Brennan et al 1987, Hudson et al 1986, Gordon et al 1993):
  - SUBJ < OBJ < OTHERS
  - A student was here to see John today: A STUDENT < JOHN</p>
- INFORMATION STATUS (Strube and Hahn, 1999):
  - HEARER-OLD < MEDIATED < HEARER-NEW</p>
  - A student was here to see John today: JOHN < A STUDENT</p>
- THEMATIC ROLES (Cote, 1998)
- FIRST MENTION / LINEAR ORDER (Rambow, 1993; Gordon et al, 1993)
  - In Lisa's opinion, John shouldn't have done that

**Massimo Poesio:** 

Skip for the moment

Eventually should also add discusson of Greene McKoon and Ratcliff, which seem to have proposed a strong version of Rule 1 themselves

## Variants of the claim

- Different definitions of CB: Grosz et al 1983, Gordon et al 1993, Passonneau 1993
- Different versions of Rule 1: Greene McKoon and Ratcliff 1992, Gordon et al 1993
- Different definitions of Rule 2:
  - Brennan et al, 1987
  - Strube and Hahn, 1999
  - Kibble, 2001

#### **Methods**

- A systematic comparison only possible by computational means, using a corpus
  - The GNOME corpus (Poesio, 2000)
  - Reliably annotated
- Problem: how to annotate a 'parametric' theory?
  - Solution: mark up the 'building blocks': sentences, (clausal) units, NPs, anaphoric information
- (Perl) Scripts use the annotation to simulate local focus construction (compute utterances, CFs, and CB) and test the claims

#### The GNOME corpus

 Initiated at the University of Edinburgh, HCRC / continued at the University of Essex

#### • 3 Genres (about 3000 NPs in each genre)

- Descriptions of museum pages (including the ILEX/SOLE corpus)
- ICONOCLAST corpus (500 pharmaceutical leaflets)
- Tutorial dialogues from the SHERLOCK corpus

#### **Massimo Poesio:**

In addition to the psychological techniques, our work in GNOME has involved a lot of corpus studies.

### An example GNOME text

#### **Cabinet on Stand**

The decoration on this monumental cabinet refers to the French king Louis XIV's military victories. A panel of marquetry showing the cockerel of France standing triumphant over both the eagle of the Holy Roman Empire and the lion of Spain and the Spanish Netherlands decorates the central door. On the drawer above the door, gilt-bronze military trophies flank a medallion portrait of Louis XIV. In the Dutch Wars of 1672 - 1678, France fought simultaneously against the Dutch, Spanish, and Imperial armies, defeating them all. This cabinet celebrates the Treaty of Nijmegen, which concluded the war. Two large figures from Greek mythology, Hercules and Hippolyta, Queen of the Amazons, representatives of strength and bravery in war, appear to support the cabinet.

The fleurs-de-lis on the top two drawers indicate that the cabinet was made for Louis XIV. As it does not appear in inventories of his possessions, it may have served as a royal gift. The Sun King's portrait appears twice on this work. The bronze medallion above the central door was cast from a medal struck in 1661 which shows the king at the age of twenty-one. Another medallion inside shows him a few years later.

## Information marked up in the GNOME corpus

- Syntactic features: grammatical function, agreement
- Semantic features:
  - Logical form type (term / quantifier / predicate)
  - `Structure': Mass / count, Atom / Set
  - Ontological status: abstract / concrete, animate
  - Genericity
  - 'Semantic' uniqueness (Loebner, 1985)
- Discourse features:
  - Deixis
  - Familiarity (discourse new / inferrable / discourse old) (using anaphoric annotation)
- A number of extra features automatically computed (e.g., is an entity the current CB, if any)

# The GNOME annotation of NEs in XML format

<ne id="ne109" cat="this-np" per="per3" num="sing" gen="neut" gf="np-mod" Iftype="term" onto="concrete" ani="inanimate" structure="atom" count="count-yes" generic="generic-no"deix="deix-yes" reference="direct" loeb="disc-function" > this monumental cabinet </ne>

# The MATE/GNOME markup scheme for anaphoric information

<NE ID="ne07">Scottish-born, Canadian based jeweller, Alison Bailey-Smith</NE> <NE ID="ne08"> <NE ID="ne09">Her</NE> materials</NE>

<ANTE CURRENT="ne09" REL="ident"> <ANCHOR ANTECEDENT="ne07" /> </ANTE>
### Reliability

- Crucial requirement for the corpus to be of any use, is to make sure that annotation is RELIABLE (I.e., two different annotators are likely to mark in the same way)
- E.g., make sure they can agree on part-of-speech tag - ... we walk in SNAKING lines (JJ? VBG?)
- Or on attachment
- Agreement more difficult the more complex the judgments asked of the annotators
  - E.g., on givenness status
- Often a detailed ANNOTATION MANUAL required
- Task must also have to be simplified

# A measure of agreement: the K statistic

- Carletta, 1996: in order for the statistics extracted from an annotation to be reproducible, it is crucial to ensure that the coding distinctions are understandable to someone other than the person who developed the scheme
- Simply measuring the percentage of agreement does not take chance agreement into account
- The K statistic (Siegel and Castellan, 1988):
  - K=0: no agreement
  - .6 <= K < .8: tentative agreement
  - .8 <= K <= 1: OK agreement

#### The GNOME annotation manual: Markables

- ONLY ANAPHORIC RELATIONS BETWEEN NPs
- DETAILED INSTRUCTIONS FOR MARKABLES
  - ALL NPs are treated as markables, including predicative NPs and expletives (use attributes to identify non-referring expressions)

# Achieving agreement (but not completeness) in GNOME

- RESTRICTING THE NUMBER OF RELATIONS
  - IDENT (John ... he, the car ... the vehicle)
  - ELEMENT (Three boys ... one (of them))
  - SUBSET (The vases ... two (of them) ... )
  - Generalized POSSession (*the car ... the engine*)
  - OTHER (when no other connection with previous unit)

### Limiting the amount of work

- Restrict the extent of the annotation:
  - ALWAYS MARK AT LEAST ONE ANTECEDENT FOR EACH EXPRESSION THAT IS ANAPHORIC IN SOME SENSE, BUT NO MORE THAN ONE IDENT AND ONE BRIDGE;
  - ALWAYS MARK THE RELATION WITH THE CLOSEST PREVIOUS ANTECEDENT OF EACH TYPE;
  - ALWAYS MARK AN IDENTITY RELATION IF THERE IS ONE; BUT MARK AT MOST ONE BRIDGING RELATION

#### **Agreement results**

- RESULTS (2 annotators, anaphoric relations for 200 NPs)
  - Only 4.8% disagreements
  - But 73.17% of relations marked by only one annotator
- The GNOME annotation scheme:
  - <u>http://www.hcrc.ed.ac.uk/~poesio/GNOME/anno\_ma</u> <u>nual\_4.html</u>

## Ambiguous anaphoric expressions in the MATE/GNOME scheme

3.3: <NE ID="ne01">engine E2</NE> to

<NE ID="ne02">the boxcar at ... Elmira</NE>

5.1: and send <NE ID="ne03">*it*</NE> to <NE ID="ne04">*Corning*</NE>

<ANTE CURRENT="ne03" REL="ident"> <ANCHOR ANTECEDENT="ne01" /> <ANCHOR ANTECEDENT="ne02" /> </ANTE>



### The scripts

- A Perl script computes utterances, their CFs, and their CB on the basis of the current instantiation, and tests the validity of the claims
- Some of the parameters:
  - CBDef (which def of CB to use)
  - Utterance-related: uttdef, prevutt
  - Realization-related: direct / indirect, CF filter (e.g., exclude second person pronouns), trace behavior
  - Ranking: gf, linear order, gf+linear order, Strube-Hahn, ...
- Can be run on-line:

http://cswww.essex.ac.uk/staff/poesio/cbc

### **The Vanilla Instantiation**

- CBdef =Grosz et al 1995
- Segmentation=Walker 1989
- Utterance=finite clause
- Prev-utterance=Kameyama
- Realization=direct
- Second person pronouns do not introduce CFs
- Ranking=Grammatical function

## Utterance and CF statistics with the Vanilla instantiation

Number of utterances	1007
(of which segment boundaries)	225
Number of CFs	3039

#### **Constraint 1**

	$CF(U_i) \cap CF(V_i)$	(U <sub>i-1</sub> ) ≠ Ø	357 (3	85.4%)	]
	Have exactly	one CB	346 (3	84.4%)	
	More than or	ne CB	11 (1%	%)	
	Without a CE segment bou	8, but ndary	163 (1	6.2%)	_
	No CB		487 (4	8.4%)	
S	Strong C1	+346 (+509)	, -498	p > .9 (sig	gn test)
V	Veak C1	+833, -11		p < .001	

# An example of violation of (Strong) C1

(u1) <u>These "egg vases"</u> are of exceptional quality:

(u2) basketwork bases support <u>egg-shaped</u> bodies

(u3) and bundles of straw form the handles,

(u4) while small eggs resting in straw nests serve as the finial for <u>each lid</u>.

(u5) Each vase is decorated with inlaid decoration:

### Rule 2 (BFP): Statistics

Continuation	70 (6.9%)
Retain	38 (3.8%)
Smooth Shift	37 (3.7%)
Rough Shift	23 (2.3%)
Establishment	189 (18.8%)
Total continuous	334 (33.2%)
Zero	168 (16.7%)
Null	482 (47.9%)

#### **Rule 2: evaluation**

- The Rule is verified by a Page Rank test (p < .01)
- However, only 15% of utterances fall into one of the categories considered by the Rule
  - And 'continuing' transitions (CON + RET + SSH + RSH + EST) about 35%
- Almost half of the utterances are NULL transitions
- Other versions of the Rule (Grosz et al, Strube and Hahn, Kibble) perform much worse

#### **Rule 1: statistics**

Total number of R1-pronouns	287
3 <sup>rd</sup> Person Pronouns	
Rel. pronouns & complementizers	70

Total number of CB realizations	374
of which realized as R1-pronouns	206 (55%)
CBs NOT realized as R1-pronouns	168 (45%)
R1-pronouns not realizing CBs	81 (28.2%)

#### **Rule 1: Results**

- Not all mentions of CBs are pronominalized.
  - Rule 1 "Gordon et al" is not verified by any of the instantiations we consider
- Rule 1 (GJW 95) clearly verified: + 265 (96.7%), -9 (3.3%)
  - Two types of violations: intrasentential pronouns, pronouns referring to `global topic'
- Not all R1-pronouns refer to CBs
  - Two cases: utterances with no CB, utterances in which CB already pronominalized

# An example of violation of Rule 1 (GJW 95)

(u1) A child of 4 years needs about a third of the adult amount;

(u2) A course of treatment for a child should not normally last more than five days

(u3) unless your doctor has told you to use <u>it</u> for longer.

# Other ways of setting the parameters

- Utterances
- Realization
- Ranking
- (for more, try the website!)

## Effect of changes in the utterance parameters on Strong C1



## Effect of changes in the utterance parameters on Strong C1



# Effect of changes in utterance parameters on Rule 1 (GJW 95)



#### **Effect on Rule 2**

- Small changes, but the rule is verified by all variants
- Effect of identifying utterances with sentences:
  - Increase in 'Continuous' transitions (CON+ RET + SSH + RSH + EST) (from 35.4% to 43%)
  - large increase in Rough Shifts (more RSH than SSH) and Retains

#### **Effect of the utterance parameters**

- Reducing the number of utterances leads to a reduction in the violations to Strong C1
- Best results: identifying utterances with sentences (u=s)
- Allowing update after every clause (also non finite) only acceptable when assuming traces
- With u=f, 'Suri' definition of previous utterance significantly better than Kameyama's, but improvement minor
- There is a tradeoff between Strong C1 on one side, and Rule 1 and Rule 2, on the other

## The tradeoff between Constraint 1 and Rule 1



# Changing the Realization parameters

- The least studied parameter
- Possibilities:
  - Indirect realization
  - Second person pronouns as CF realizations
  - Other candidates for realization (not discussed): predicative NPs; implicit anaphors in bridging references;
- Considering both finite clauses and sentences as possible utterances

### Constraint 1 (indirect realization, u=f)

$CF(U_{i}) \cap CF(U_{i-1}) \neq \emptyset$	547 (56.3%)
Have exactly one CB	525 (54%)
More than one CB	22 (2.3%)
Without a CB, but segment boundary	122 (12.6%)
No CB	303 (31.2%)

Strong C1

+525, -325

p < .

p < .001 (sign test)

### Constraint 1 (indirect realization, u=s)

$CF(U_{i}) \cap CF(U_{i-1}) \neq \emptyset$	416 (62.2%)
Have exactly one CB	390 (58.3%)
More than one CB	26 (3.9%)
Without a CB, but segment boundary	102 (15.2%)
No CB	151 (22.6%)

Strong C1 +390 (+492), -177 p < .001 (sign test)

## Effect of parameter changes on Strong C1 and Rule 1



## A violation to Rule 1 caused by indirect realization

(u39) *The two stands* are of the same date as the coffers, but were originally designed to hold rectangular cabinets.

(u42) One stand was adapted in the late 1700s or early 1800s to make it the same eight as the other.

### Rule 2 (BFP): Statistics (u=f)

Continuation	89 (9.2%)
Retain	127 (13.1%)
Smooth Shift	64 (6.6%)
Rough Shift	97 (10%)
Establishment	170 (17.5%)
Total continuous	547 (54.3%)
Zero	138 (14.2%)
Null	287 (29.5%)

### Rule 2 (BFP): Statistics (u=s)

Continuation	63 (9.4%)
Retain	114 (17%)
Smooth Shift	35 (5.2 %)
Rough Shift	100 (14.9%)
Establishment	104 (15.5%)
Total continuous	416 (62.2%)
Zero	98 (14.6%)
Null	155 (23.2%)

### **Effect of the realization parameters**

- Allowing for indirect realization is sufficient to satisfy Strong C1, even with u=f
- Negative effects on Rule 1, but especially Rule 2 (not verified)
  - 50% more RET than CON, and more RSH than SSH
- Treating 2per pronouns as CF realizations improves results, but theoretically dubious

#### **Ranking parameters**

- In addition to grammatical function, considered
  - Combination of GF and linear order
  - Linear order
  - Information status (Strube and Hahn, 1999)
- Also considered ranking for complex NPs (Gordon et al, 1999; Poesio and Nissim, 2001)

#### **GFTHERELIN**

- Grammatical function + linear order to disambiguate + there-obj ranked as subjects
- Eliminates Multi-CB utterances (violations of Weak Constraint 1)

# The GFTHERELIN instantiation at-a-glance

Instantiation	Strong C1	Rule 1	Rule 2
DF	+352, -450 (46.3%)	+291, -11 (3.6%)	.001
DS	+273, -249 (37.2%)	+259, -14 (5.1%)	.001
IF	+529, -310 (31.9%)	+463,-18 (3.7%)	1
IS	+408, -157 (23.5%)	+378, -30 (7.4%)	.05
#### Linear order

- No changes on Strong C1 (as expected)
- Same results for Rule 1
- Slightly worse results for Rule 2

# Strube and Hahn, 1999

- HEARER-OLD < INFERRABLE < HEARER-NEW
- No differences with Strong C1 and Rule 1 (although different violations)
- Rule 2: verified at the .01 level by both IF and IS instantiations
- Rule 2 (Strube and Hahn): not verified.

# A violation to Rule 1 with Strube-Hahn, but not GFTHERELIN

(s67) An inventory of Count Branicki's possessions made at his death describes both the corner cupboard and the objects displayed on its shelves: a collection of mounted Chinese porcelain and clocks, some embellished with porcelain flowers.

(s68) The drawing of the corner cupboard, or more probably an engraving of <u>it</u>, must have caught Branicki's attention.

# A violation to Rule 1 with GFTHERELIN, but not Strube-Hahn

(s88) Christened by his contemporaries as 'the most skilled artisan in Paris', <u>Andre-Charles</u> <u>Boulle's name</u> is synonymous with *the practice of veneering furniture with marquetry of tortoiseshell, pewter, and brass.* 

(s89) Although he did not invent *the technique*, Boulle was *its* greatest practitioner and lent <u>his</u> <u>name</u> to *its* common name: Boulle work.

# Rule 2 (BFP): Statistics (Strube-Hahn, u=s)

Continuation	94 (14.1%)	
Retain	103 (15.4%)	
Smooth Shift	44 (6.6%)	
Rough Shift	60 (9.0%)	
Establishment	107 (16%)	
Total continuous	408 (61%)	
Zero	101 (15.1%)	
Null	160 (23.9%)	

# **Ranking variants: a summary**

- GFTHERELIN eliminates the Multiple-CB cases
- GFTHERELIN and linear order equivalent
- Strube-Hahn reduces violations to Rule 2

# Discussion, I: Comparing Instantiations

- 'Vanilla' instantiation: only Rule 1 (GJW 95, 83) convincingly verified
- Allowing indirect realization, or identifying utterances with sentences, results in Strong C1 being verified as well
- But thes changes also result in the number of violations to Rule 1 doubling, and in Rule 2 not being verified

# **Beyond Entity Coherence: Relations**

(u1) If it will not stick on again, use a new one.

(u2) THEN, continue your previous treatment schedule as normal.

# Beyond Entity Coherence (Oberlander & Poesio, in progress)

(u1) On the drawer above the door, gilt-bronze military trophies flank a medallion portrait of Louis XIV.

(u2) In the Dutch Wars of 1672-1678, France fought simultaneously against the Dutch, Spanish, and Imperial armies, defeating them all.

(u3) This cabinet celebrates the Treaty of Nijmegen, which concluded the war.

# A Hybrid view of coherence

 Entity continuity only one among the preferences holding at the discourse level (which would also include constraints specifying preferences for rhetorical and temporal coherence)

#### **CB Uniqueness?**

(s68) The drawing of <u>the corner cupboard</u>, or more probably an engraving of it, must have caught *Branicki*'s attention.

(s69) Dubois was commissioned through a Warsaw dealer to construct <u>the cabinet</u> for *the Polish aristocrat*.

# Variety

- Ensuring variety seems to be as important a preference as maintaining coherence:
  - Only about half of CBs realizes as pronouns
  - CBs very seldom maintained for more than twothree utterances
  - Number of retains comparable to number of continues

# **Demonstratives and local salience**

• The Poesio / Modjeska work

#### **Global salience**

#### Massimo Poesio:

Main point: antecedent much faster when in context 3 than in 2 or 1

# **Arguments for globa** Clark and Sengul's e

In fact, this has more to do with local focus / global focus separation than propositional vs. situational (perhaps Lesgold 1972 would have been better!)

Harald: but did they control for position of antecedent in sentence? Subjects read a context with t Animacy/inanimacy? (psych student: usually you just permutate same

target sentence containing a r mentioned in context sentence sentences) noun)

• Example:

The campers erected a shelter of lightweight plastic in the clearing.

They piled a supply of dry wood next to a tree.

The smell of coffee mingled with smoke from the campfire.

TARGET2: <u>The tree</u> was an enormous redwood.

# **Clark and Sengul's experiments**

#### • Results:

Type of nominal	Context Sentenc e 1	Context Sentenc e 2	Context Sentenc e 3
Full NP	2174	2166	1802
Pronoun	2280	2133	1847
Mean	2227	2150	1825

• Significant effect of distance: F'=18.99, p < .001

# The Anderson et al study

