Proceedings of the

# Fifteenth Amsterdam Colloquium

December 19 — 21, 2005

# Proceedings of the Fifteenth Amsterdam Colloquium

December 19 — 21, 2005

Paul Dekker and Michael Franke (eds.)

ILLC/Department of Philosophy
University of Amsterdam

Printed by Grafisch Centrum Amsterdam
Cover design by Crasborn Grafisch Ontwerpers bno
ISBN: 90-5776 1467

## Preface

The 2005 edition of the Amsterdam Colloquium is the Fifteenth in a series which started in 1976. Originally, the Amsterdam Colloquium was an initiative of the Department of Philosophy of the University of Amsterdam. Since 1984 the Colloquium is organized by the Institute for Logic, Language and Computation (ILLC) of the University of Amsterdam.

These proceedings contain the abstracts of the papers presented at the colloquium. In the first section one can find abstracts of the talks given by some of the invited speakers, Sigrid Beck and Nissim Francez (joint talk with Gilad Ben-Avi). The next two sections contain contributions to the two workshops:

≫ Language and Learning
≫ Semantic Universals

The fourth section consists of the contributions to the general program. In all cases the copyright resides with the individual authors.

For the organization of the Fifteenth Amsterdam Colloquium financial support is received from:

- the Royal Dutch Academy of Sciences (KNAW)
- the Netherlands Organization for Scientific Research (NWO)
- the Institute for Logic, Language and Computation (ILLC)
- the NWO-funded project 'Formal Language Games'
- Springer
- the city of Amsterdam

which is gratefully acknowledged.

The organizers would like to thank the authors for their contribution and of course the members of the program committee for the great job they have done:

≫ (local committee) Johan van Benthem, Martin Stokhof (chair), Henk Zeevat
≫ (the invited speakers) Sigrid Beck, Nissim Francez, Manfred Krifka, Lawrence S. Moss
≫ (external committee) David Beaver, Bart Geurts, Jack Hoeksema, Marcus Kracht, Angelika Kratzer, Michael Moortgat, Henriette de Swart, Ede Zimmermann

*The Editors*
*Amsterdam, November 2005*

**Contents**

# Invited Speakers

# A SECOND TIME AND AGAIN

SIGRID BECK

Englisches Seminar
Universität Tübingen
Sigrid.Beck@uni-tuebingen.de

This paper considers focus alternatives to presuppositional elements like *again*. We observe that there are empirical differences between *again* and its non-presuppositional counterpart *a second time*. A general question is raised about presuppositions in alternative sets.

## 1. Introduction

It has been observed that the discourse behaviour of focused *again* differs from that of unfocused *again* (Fabricius-Hansen 1983, Kamp & Rossdeutscher 1994, among others). An example taken from Beck (to appear) is given in (1) (imagine somebody reading through a long list of former US presidents).

(1) a. Smith was a Republican, Jones was a Republican,
Longbottom was a Republican AGAIN

b. Smith was a Republican, Jones was not a Republican,
Longbottom was a Republican again/ *AGAIN

Recent discussion of several interesting aspects of this problem is found e.g. in Klein 2001 and Beck (to appear). Here I will simply raise the question of the focus semantic contribution of a presuppositional element like *again*. Observe that *again* does not license the same contrast relationships as the non-presuppositional, but otherwise semantically parallel *a second time/for the second time*. Hence *again* and *a second time* must introduce different focus alternatives.

(2) a. ?? Peter is in Rome for the first time & Paul (is) AGAIN.

b. Peter is in Rome for the first time & Paul (is) for the SECOND time.

Section 2 discusses the effect of focus on *again* in more detail, and introduces a second purely presuppositional element, *also*. Section 3 generalizes the question about presupposition (ff: psp) in alternative sets. Conclusions are drawn in section 4.

## 2. Focus on Purely Presuppositional Items

### 2.1. Again

I will work with the (simplified) semantics in (3) for *again*. I suggest that typical focus alternatives (ff: FAlts) to *again* are a semantically empty adverb (i.e. the identity function of the relevant type) and *still*. There may be further plausible FAlts like *(not) yet* and *(not) anymore* (thanks to Graham Katz for pointing this out). It is also likely that the set of FAlts varies with context. I will concentrate on the semantically empty adverb and *still*. Let $ALT_x$ be the set of contextually relevant focus alternatives to expression *x*. This set will contain the focused element itself as well as its alternatives. The assumption that we have a typical set of alternatives $ALT_{again}$={[[again]], [[still]], ∅} explains discourse coherence in the exchanges below. Here I use contrast to test whether something is an FAlt to *again* or *still*. A category □ stands in a contrast relation to a category □ if the the ordinary semantic value of □ is a member of the focus semantic value of □, i.e. $[[\square]]_o \in [[\square]]_f$, and $[[\square]]_o \neq [[\square]]_o$ (Rooth (1992a)). Regarding (4A-4Bb), for example, a natural analysis in the framework of Rooth (1992a) would be to regard (4A) as the focus antecedent for (4Bb), as indicated in (6). This implies (6c), which in turn implies that an FAlt to *again* is the empty adverb.

(3) [[again]] (p<□,<s,t>>) (t) (w) = 1 if p(t)(w) & □t'[t'<t & p(t')(w)]
= 0 if ~p(t)(w) & □t'[t'<t & p(t')(w)]
undefined otherwise.

(4) **A**: Ellen is the president.
**B**: a. (Yes,) Ellen is STILL the president.
b. (Yes,) Ellen is the president AGAIN.

(5) **A**: Ellen is still the president. **B**: Ellen is the president AGAIN.

(6) a. $g(C) \in [[\square]]_f^g$ & $g(C) \neq [[\square]]_o^g$
b. [[Ellen is the president AGAIN] ~C]
g(C) := $[[\text{Ellen is the president}]]_o$
c. $[[\text{Ellen is the president}]]_o \in [[\text{Ellen is the president AGAIN}]]_f$

(7a,b) are examples of sentence internal contrast that show the same. I assume that in cases in which a sentence S2 contains an ellipsis that finds its antecedent in a sentence S1, S2 must stand in a contrast relation to S1 - i.e. $[[S1]]_o \in [[S2]]_f$ (Rooth (1992b)); the account can be extended to deaccenting (7b). (8a) is an example in which 'be in Rome' and 'be in Rome again' are scalar FAlts. (8b) may be an example of association with scalar *only*. Both have a metalinguistic flavour because we try to focus a psp, but

they are not unacceptable. It is instructive to contrast the data above with examples that do not work. *Again* being an adverb that combines with a proposition to yield a proposition, perhaps other adverbs of the same type could be FAlts, too? This is not generally plausible, as (9) illustrate. Hence the tests I ran above are meaningful tests and do indeed tell us something about FAlts to *again*.

(7) a. Peter is still in Rome and Paul is AGAIN.
b. Peter is in Rome and Paul is in Rome AGAIN.

(8) a. Peter is in Rome. He is even in Rome AGAIN.
b. Peter is only in Rome AGAIN - he is not STILL in Rome.
c. scale: 'be in Rome' < 'be in Rome again' < 'still be in Rome'

(9) a. ?? Peter is probably/often in Rome and Paul is AGAIN.
b. ?? Peter is only PROBABLY in Rome - he is not STILL in Rome/ in Rome AGAIN.

A more minimal contrast exists between *again* and *for the second time* concerning their respective FAlts (thanks to Irene Heim (p.c.) for pointing this out). I will assume the lexical entry for *for the second time* given in (10). The contrast between (11a) and (11b) shows that it matters for the purpose of FAlts whether a meaning component is asserted or presupposed: what is presupposed by *again* -- $\exists t'[t'<t \ \& \ p(t')(w)]$ -- is asserted by *for the second time*. *For the first time* is an FAlt to *for the second time*, but not to *again*. Given these observations, I suggest the hypothesis in (12) (a purely presuppositional element is one that, like *again*, triggers a psp but has no effect on the assertion).

(10) [[for the second time]] $(p<\exists,<s,t>>)\ (t)\ (w) = 1$ iff $p(t)(w) \ \& \ \exists t'[t'<t \ \& \ p(t')(w)]$

(11) a. ?? Peter is in Rome for the first time & Paul (is) AGAIN.
b. Peter is in Rome for the first time & Paul (is) for the SECOND time.

(12) **Hypothesis**: Focus alternatives to purely presuppositional items are other purely presuppositional items plus the empty alternative of the same type.

There remains the larger question of how psps show up in FAlts, and whether we can predict the facts we just observed about *again* systematically from the answer to that question. Below I will take a look at another purely presuppositional element, and then I will briefly comment on the more general question.

### 2.2. Also

Another purely presuppositional element is *too*/*also*. We can simplify and assume (14) about its semantic contribution. The application is illustrated in (15) where we suppose that the associate of *also* is *Bill*.

(14) [[also]] $(y)(P)(t)(w)$ $=1$ if $P(y)(t)(w) \ \& \ \exists x[x \neq y \ \& \ P(x)(t)(w)]$
$=0$ if $\sim P(y)(t)(w) \ \& \ \exists x[x \neq y \ \& \ P(x)(t)(w)]$

undefined otherwise.

(15) Bill also ran. [[also]] (Bill) ([[run]]) (**t1**)(w)
=1 if ∃x[x≠Bill & x ran in w at **t1**] & Bill ran in w at **t1**
=0 if ∃x[x≠Bill & x ran in w at **t1**] & ~(Bill ran in w at **t1**)
undefined otherwise.

We may ask what *also*'s FAlts are, and whether it is plausible that an empty element is among them. I think this is so. Below is an example that illustrates that. B's utterance is the natural focus antecedent for utterance A2. The assumption that an FAlt to *also* is the empty alternative helps with the data in (17): we can say that there is a contrast relation between the first and the second sentence in the conjunction. This reduces (17a) to the same phenomenon as (17b,c). I should note that in the case of FAlts to *too*/*also*, there is in fact some discussion in the literature, although it is not very prominent. Krifka (1999) and following him Rullmann (2003) are concerned with focused *too*/*also*. They propose that the FAlt to *too* is negation. The reasoning is that *too* (ignoring its psp) expresses the identity function on propositions, and the only relevant alternative to that is negation. Dimroth (2004) argues on empirical grounds that German *auch* 'also' contrasts with verum focus, stressed negation, and the affirmative particles *schon*/*wohl*. Assuming that verum focus contributes the relevant identity function, Dimroth's position agrees with the suggestions made here.

(16) **A1**: Es ist kalt. **B**: Es regnet.
It's cold. It is raining
**A2**: Regnen tut's AUCH.
rain does it too
'It's raining, too'

(17) a. Peter is in Rome and PAUL is, TOO.
b. Peter will see Jim and PAUL will STACEY.
c. Peter saw Jim and PAUL DIDN'T.

## 3. Focus Alternatives to Presuppositional Elements

How do psps fare in general when alternative sets are constructed? Intuitively, FAlts to a given element are things that are comparable but contrasting. Normally, we assume that comparable are things that have the same type (Cohen (1999) has a more constrained view of what plausible FAlts are, but he is not concerned with classical psps). Looking at presuppositional elements reveals that information on semantic type is not enough to tell us what is comparable to a given focused item. *Again* and *for the first time* (also *often* and *probably*) have the same type, but that is not sufficient to make them FAlts. Recall that I introduced $ALT_x$ - the set of relevant FAlts to an element *x*.

We want to predict what those can be in a systematic way. It seems to me that it is required that the psps of the elements of $ALT_x$ be parallel, in some sense, to the psps of *x*. To start with fairly obvious cases, note that items with the same psps or no psp at all are FAlts (18); and note also that items whose psps are unrelated, or presuppositional vis-a-vis non-presuppositional items are not FAlts (19). The picture is complicated by items that have psps that seem related, but not identical. For example, *still* and *again* can be focus alternatives, while *again* and *for the first time* cannot. This must be because *still* and *again* share a psp about a preceding time interval. Similarly for *start - stop* vs. *start - try*, *start - manage* as illustrated below. We are looking for a definition of the general shape of (23) which will permit us to predict what the FAlts to a given expression can be. We still need to define what it means for psps to be parallel.

(18) **plausible alternatives**:

| a. | no (relevant) psp: | b: | same psps: |
|---|---|---|---|
| | *sing - dance* | | *both - neither* |

(19) **non-alternative**s:

| a. | psp vs. no psp: | b. | non-parallel psps: |
|---|---|---|---|
| | *for the first time - again* | | *start - manage* |
| | *start - try* | | |

(20) [[still]] (p)(t) =1 iff p(t) & ∃t'[p(t') & t' extends to t]
=0 iff ~p(t) & ∃t'[p(t') & t' extends to t]
undefined otherwise.

(21) a. Molly started to play soccer and Sue stopped _ .
b. ?? Molly started to play soccer and Sue tried _ .
c. ?? Molly started to play soccer and Sue managed _ .

(22) **"parallel" psps - plausible alternatives**:
*again - still;* *stop - start*

(23) For any expression x of type τ, the set of plausible focus alternatives to x,
$ALT_x$={y:y is of type τ and the psps of x ARE PARALLEL TO the psps of y}

I leave this as a project for future research. A final comment: there is an interesting relationship between the issue discussed here and Abusch's (2002) suggestion to derive certain psps from the alternatives that lexical items give rise to. For example, the psp of "x be right that p" that x believes p would arise because the alternative to "be right" is "be wrong", which shares this meaning component, and there is a pragmatic psp that some alternative is true. We approach the problem from opposite perspectives, in that Abusch wants to predict psps from alternatives, while I want to predict FAlts from psps. That is, I would wish to predict that an FAlt to "x be right hat p" is "x be wrong that p" from the fact that they have the same psp that x believes p. Interestingly, though, the

same issue concerning "parallel" psps in cases like *again* arises under both strategies.

## 4. Conclusion

Presuppositions matter for what an element's focus alternatives are. Items that share psps are alternatives. There is a little more leeway: *still* and *again* are alternatives though their psps aren't exactly the same. There is also the special case that an alternative to a purely presuppositional element is the identity function of the same type. Items with unrelated psps are not focus alternatives, though, A precise definition is missing of when psps are sufficiently alike.

**Acknowledgements**

I would like to thank Irene Heim, Graham Katz, Mats Rooth, and audiences at U. Potsdam, UConn, UMass, U. Göttingen, U. Tübingen, the Milan Meeting 2004, and the Workshop on Information Structure (2004, Universität Stuttgart.

**Bibliography**

Abusch, D. 2002. Lexical Alternatives as a Source of Pragmatic Presuppositions. SALT12.

Beck, S. (to appear). Focus on *Again*. To appear in: Linguistics and Philosophy.

Cohen, A. (1999). How are Alternatives Computed? *Journal of Semantics* 16, 43-65.

Dimroth, C. (2004). Fokuspartikeln und Informationsgliederung im Deutschen. Studien zur Deutschen Grammatik 69, Stauffenburg, Tuebingen.

Fabricius-Hansen, C. (1983). Wieder ein *wieder*? Zur Semantik von *wieder*. In: R. Baeuerle et al. (eds.): *Meaning, Use and Interpretation of Language*. De Gruyter, Berlin, 97-120.

Kamp, Hans and Antje Rossdeutscher (1994). DRS-construction and lexically-driven inference. *Theoretical Linguistics* 20: 165-235.

Klein, W. (2001). Time and again. In: C. Fery & W. Sternefeld: *Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow.* Akademie, Berlin, 267-286.

Krifka, M. (1999). Additive Particles under Stress. Proceedings of SALT 8, 111-128.

Rooth, M. (1992a). A Theory of Focus Interpretation. *NaLS* 1, 75-116.

Rooth, M. (1992b). Ellipsis Redundancy and Reduction Redundancy. In: Berman, S. & A. Hestvik (eds.): Proceedings of the Stuttgart Ellipsis Workshop.

Rullmann, H. (2003). Additive Particles and Polarity. *Journal of Semantics* 20,329-401.

# PROOF-THEORETIC SEMANTICS FOR A SYLLOGISTIC FRAGMENT

GILAD BEN AVI AND NISSIM FRANCEZ

Computer Science dept.,
Technion-IIT, Haifa, Israel
bagilad,francez@cs.technion.ac.il

## 1. Introduction

We present some *prolegomena* to *Proof-Theoretic Semantics (PTS)* for natural language ($NL$). The following quotation from Schroeder-Heister 2005 emphasizes the lack of applicability to $NL$, the original reason for $PTS$ to start with:

> Although the *"meaning as use"* approach has been quite prominent for half a century now and provided one of the cornerstones of philosophy of language, in particular of ordinary language philosophy, it has never become prevailing in the *formal* semantics of artificial and natural languages. In formal semantics, the *denotational* approach which starts with interpretations of singular terms and predicates, then fixes the meaning of sentences in terms of truth conditions, and finally defines logical consequence as truth preservation under all interpretations, has always dominated.

In order to device a PTS for (a fragment of) NL, two steps are required:
1. Device a *proof-theory (a calculus)* for the fragment, satisfying criteria proposed for PTS in logic. Replace *truth condition* by *derivability conditions* (in the above calculus) as the meaning of sentences in the fragment.
2. Identify the contribution of subsentential phrases (down to words) to the PTS meaning of sentences in which they occur.
Here, we focus on the first task only.

The studied fragment is $SYL$ (syllogistic logic) Moss 2005, where Moss considers its Hilbert-like axiomatization, being concerned mainly with completeness w.r.t. set-based "natural" semantics, and extensions not expressible in 1st-order logic. The fragment is: $All\ X\ are\ Y \quad Some\ X\ are\ Y \quad No\ X\ are\ Y \quad J\ is\ an\ X$
$X, Y$ range over predicate symbols, and $J$ as an individual constant. Here we only study the *positive* fragment $SYL^{+}$, without $No\ X\ are\ Y$.

We propose a *natural deduction* proof system for $SYL$, with proof-terms embodying a *Curry-Howard (CH)* correspondence. The system is shown to be *harmonious*, taken here as the requirement that its rules satisfy *local soundness (LS)* and *local completeness (LC)* Pfenning and Davies 2001. LS requires that *every* introduction immediately followed by elimination is *reducible* to a derivation without such *detour*. Failing LS means elimination is too strong. LC requires that for every elimination *there is* a reconstructing introduction. Failing LC means elimination is too weak.

The proof-terms are *drawn from* the traditional $\lambda$-calculus, but receive a somewhat different interpretation via a *BHK*-like justification of the deduction-rules. All rules in Moss 2005 are derivable in our system, rendering it complete w.r.t. the same evaluative semantics, though this is of no central interest here.

## 2. The natural deduction system

**A BHK-like justification**:
– A proof of $All\ X\ are\ Y$ is a (construction for) a *function* mapping a proof of $J\ is\ an\ X$ to a proof of $J\ is\ a\ Y$.

This is different from, though related to, the function involved in the *BHK*-justification for $\forall x.\phi(x)$, mapping an object $o$ to a proof of $\phi(o)$.

– A proof of $Some\ X\ are\ Y$ is a pair of proofs of $J\ is\ an\ X$ and $J\ is\ a\ Y$.

This is reminiscent to the *BHK*-justification of *conjunction*, also constituting a pair of proofs.

**The natural deduction rules**
There is an *introduction-rule* and *elimination-rule* for each kind of propositions, presented below, together with proof-terms, to which we return later. The presentation is in Gentzen-style *ND*, using sequents.

$$S : u \vdash S : u \ \ (Ax) \quad \text{any } S \in SYL$$

$$\frac{\Gamma, [J\ is\ an\ X]_i : u \vdash J\ is\ a\ Y\ : M}{\Gamma \vdash All\ X\ are\ Y : \lambda u.M}\ (All - I_i)$$

$$\frac{\Gamma_1 \vdash All\ X\ are\ Y : M \quad \Gamma_2 \vdash J\ is\ an\ X : N}{\Gamma_1\Gamma_2 \vdash J\ is\ a\ Y : (MN)}\ (All - E)$$

$$\frac{\Gamma_1 \vdash J\ is\ an\ X : M_1 \quad \Gamma_2 \vdash J\ is\ a\ Y : M_2}{\Gamma_1\Gamma_2 \vdash Some\ X\ are\ Y : \langle M_1, M_2 \rangle}\ (Some - I)$$

$$\frac{\Gamma_1 \vdash Some\ X\ are\ Y : M \quad \Gamma_2, [J\ is\ an\ X]_i : u, [J\ is\ a\ Y]_i : v \vdash S : N}{\Gamma_1\Gamma_2 \vdash S : \ \mathbf{let}\ \langle u, v \rangle = M\ \mathbf{in}\ N}\ (Some - E_i)$$

$$\frac{\Gamma, [Some\ X\ are\ Y]_i \vdash S}{\Gamma \vdash No\ X\ are\ Y}\ (No - I)_i^S$$

Here $S$ is a *parameter proposition* not occurring in $\Gamma \cup \{Some\ X\ are\ Y\}$.

$$\frac{\Gamma_1 \vdash Some\ X\ are\ Y \quad \Gamma_2 \vdash NoXY}{\Gamma_1\Gamma_2 \vdash S}\ (No - E)$$

We denote by $\vdash_{ND-Syl}$ derivability/provability in this system. As an example of a derivation using those rules, consider $Some\, X\, are\, Y,\ All\, Y\, are\, Z \vdash_{ND-syl^+} Some\, X\, are\, Z$.

$$\dfrac{Some\ X\ are\ Y : x \qquad \dfrac{[J\ is\ an\ X]_i : u \qquad \dfrac{[J\ is\ a\ Y]_i : v \quad All\ Y\ are\ Z : w}{J\ is\ a\ Z : (wv)}\ (All-E)}{Some\ X\ are\ Z : \langle u, (wv)\rangle}\ (Some-I)}{Some\ X\ are\ Z :\ \mathbf{let}\ \langle u, v\rangle\ = x\ \mathbf{in}\ \langle u, (wv)\rangle}\ (Some-E_i)$$

## 3. Properties of the Positive Fragment

### 3.1. Curry-Howard correspondence

We point out several observations about derivations in $ND - syl^+$.

1. The conclusion of an instance of application of the $(All-I)$ rule cannot serve as a premiss of another instance of application of the same rule.
2. The conclusion of an instance of application of the $(All-E)$ rule cannot serve as a major premiss of another instance of application of the same rule.
3. The conclusion of an instance of application of the rule $(Some-I)$ cannot serve as a premiss of another instance of application of the same rule.
4. The conclusion of an instance of application of the rule $(All-I)$ cannot serve as a premiss of an instance of application of the $(Some-I)$ rule.

Two important remarks about *discharge* of assumptions by the $(All-I)$-rule:

**No vacuous discharge:** The rule $(All-I)$ should not allow *vacuous discharge*; otherwise, the following unwarranted[1] derivation becomes possible.

$$\dfrac{J\ is\ a\ Y : u \vdash J\ is\ a\ Y : u}{J\ is\ a\ Y : u \vdash All\ X\ are\ Y : \lambda v.u}\ (All - I_{vac})$$

**No multiple discharge:** In the absence of the *Weakening* structural rule, *multiple discharge* becomes actually impossible, because there is no way to generate sequents, of the form, say, $\Gamma, J\ is\ an\ X,\ J\ is\ an\ X \vdash J\ is\ a\ Y$.

However, *contraction needs to be admitted*. To see the need for it, consider the following: $J\ is\ an\ X,\ All\ X\ are\ Y,\ All\ X\ are\ Z \vdash Some\ Y\ are\ Z$.

[1] Semantically, unsound ...

The assumption $J\ is\ an\ X$ has to be used twice, to eliminate both occurrences of $All$.

$$\dfrac{\dfrac{J\ is\ an\ X : x \quad All\ X\ are\ Y : y}{J\ is\ a\ Y : (yx)}\ (All-E) \quad \dfrac{J\ is\ an\ X : x \quad All\ X\ are\ Z : z}{J\ is\ a\ Z : (zx)}\ (All-E)}{Some\ Y\ are\ Z : \langle (yx), (zx) \rangle}\ (Some-I)$$

Note the $(Some-E)$ elimination rule, that *does not* allow *projection*. Indeed, we do not want $J\ is\ an\ X$ to be derivable from $Some\ X\ are\ Y$. The reduction-rule to be shown for harmony requires a *joint discharge* of its "J-assumptions". Thus, the resulting proof-terms are "almost" *linear*. This gives rise to the definition of $\Lambda^{fl}$, the subset of the set $\Lambda$ of all $\lambda$-terms, referred to as *flat terms*.

**Definition:(flat terms)** $\Lambda^{fl}$ is the smallest subset of $\Lambda$ satisfying:

1. If $u$ is a term-variable then $u{\in}\Lambda^{fl}$.
2. If $M, N{\in}\Lambda^{fl}$ and $M$ is a variable or an abstraction-term, then $(MN){\in}\Lambda^{fl}$.
3. If $u$ is a term-variable and $M{\in}\Lambda^{fl}$ s.t. $M$ ia a variable, or an application-term containing *exactly one* free occurrence of $u$, then $\lambda u.M{\in}\Lambda^{fl}$.
4. If $M_1, M_2{\in}\Lambda^{fl}$, and none is of the form $\lambda x.N$, nor of the form $\langle P, Q\rangle$, then $\langle M_1, M_2\rangle{\in}\Lambda^{fl}$.
5. If $M, N{\in}\Lambda^{fl}$, $M \equiv \langle M_1, M_2\rangle$ or $M \equiv x$, and $N$ is a pair-term or a **let**-term, then $\mathbf{let}\ \langle u, v\rangle = M\ \mathbf{in}\ N{\in}\Lambda^{fl}$.

While the $ND-syl^{+}$ calculus uses the the flat terms as its proof-terms, a subset of the (implicational fragment of the) Intuitionistic linear propositional calculus proof terms, it constitutes a completely different *typing* system for those terms. However, it enjoys similar properties to the latter, expressed in the following two theorems.

**Theorem (flatness):** If $\vdash_{ND-syl^{+}} S : M$, then $M{\in}\Lambda^{fl}$ and $free(M) = Subjects(\Gamma)$.

**Theorem (subject construction)**: If $M{\in}\Lambda^{fl}$, then there exists a $SYL^{+}$ proposition $S$ s.t. there exists a derivation $\mathcal{D}$ of $\Gamma{\vdash}_{ND-Syl^{+}} S : M$, where:

1. If $M \equiv u$ (a term-variable), then $\Gamma = S$ for some type $S$ (a proposition in the Syllogistic fragment!), and $\mathcal{D}$ is the axiom $S : u{\vdash}_{ND-Syl^{+}} S : u$.
2. If $M \equiv (PQ)$, then the last step in $\mathcal{D}$ must be

$$\dfrac{\Gamma_1{\vdash}All\ X\ are\ Y : P \quad \Gamma_2{\vdash}J\ is\ a\ Y : Q}{\Gamma_1\Gamma_2{\vdash}J\ is\ a\ Y : (PQ)}\ (All-E)$$

   for some partition $\Gamma = \Gamma_1\Gamma_2$.

3. If $M \equiv \lambda u.N$, then the last step in $\mathcal{D}$ must be

$$\dfrac{\Gamma, [J\ is\ an\ X]_i : u \vdash J\ is\ a\ Y\ : N}{\Gamma{\vdash}All\ X\ are\ Y : \lambda u.N}\ (All-I_i)$$

4. If $M \equiv \langle M_1, M_2\rangle$, then the last step in $\mathcal{D}$ must be

$$\frac{\Gamma_1 \vdash J\ is\ an\ X : M_1 \quad \Gamma_2 \vdash J\ is\ a\ Y : M_2}{\Gamma_1\Gamma_2 \vdash Some\ X\ are\ Y : \langle M_1, M_2\rangle}\ (Some - I)$$

   for some partition $\Gamma = \Gamma_1\Gamma_2$.

5. If $M \equiv \mathbf{let}\ \langle u, v\rangle = P\ \mathbf{in}\ N$, then the last step in $\mathcal{D}$ must be

$$\frac{\Gamma_1 \vdash Some\ X\ are\ Y : P \quad \Gamma_2, [J\ is\ an\ X]_i : u, [J\ is\ a\ Y]_i : v \vdash S : N}{\Gamma_1\Gamma_2 \vdash S : \mathbf{let}\langle u, v\rangle = P\ \mathbf{in}\ N}\ (Some - E_i)$$

### 3.2. A correspondence with a sub-Intuitionistic fragment

Based on the identity of proof-terms, there is a natural isomorphism between $SYL^+$ and a fragment $ILprop$ of the implicational fragment of the Intuitionistic linear propositional calculus. Denote by $\vdash_{i-lint}$ the derivability in the standard natural-deduction proof-system for the latter (e.g., Negri 2002). Let the propositional variables in $ILprop$ be in 1-1 correspondence with the predicate variables in $SYL$. For simplicity, we just identify both sets. Define a syntactic mapping $\Pi :\ SYL^+ \Longrightarrow ILprop$ by: $\Pi(J\ is\ an\ X) = X$, $\Pi(All\ X\ are\ Y) = X \multimap Y$, $\Pi(Some\ X\ are\ Y) = X \bullet Y$ Obviously, if $\phi$ is in the range of $\Pi$, $\phi$ *has no nested implications*; also, there are no directly nested occurrences of pairing. Furthermore, abstraction-terms cannot be paired. Hence the name 'flat'. Extending $\Pi$ naturally to sets $\Gamma$, we get as a conclusion from sharing proof-terms that $\Gamma \vdash_{ND-syl^+} S : M \iff \Pi(\Gamma) \vdash_{i-lint} \Pi(\phi) : M$ (where corresponding subject variables are assumed in $\Gamma$ and $\Pi(\Gamma)$).

**A semantic digression:**

The only tautologies in $ILprop$ are of the form $X \multimap X$, reflecting the fact that the only validities in $SYL^+$ are of the form $All\ X\ are\ X$ (cf. Moss 2005).

### 3.3. Harmony

We now show that $ND - syl^+$ satisfies *harmony*, as expressed via the local soundness and completeness, (Pfenning and Davies 2001) providing *reduction* and *expansion* steps, embodying Prawitz's *inversion principle* Prawitz 1965; Prawitz 1971. For better readability, we employ the Prawitz style presentation of natural deduction.

$All\ X\ are\ Y$ **– local soundness:**

$$\dfrac{\dfrac{\overset{[J\ is\ a\ X]_i}{\overset{\mathcal{D}_1}{J\ is\ a\ Y}}}{All\ X\ are\ Y}\ (All - I_i) \qquad \overset{\mathcal{D}_2}{J\ is\ a\ X}}{J\ is\ a\ Y}\ (All - E) \quad \rightsquigarrow_r \quad \begin{array}{c}\mathcal{D}_2\\ J\ is\ a\ X\end{array} \quad \begin{array}{c}\mathcal{D}_1\\ J\ is\ a\ Y\end{array}$$

$All\ X\ are\ Y$ **– local completeness:**

$$\overset{\mathcal{D}}{All\ X\ are\ Y} \quad \rightsquigarrow_e \quad \dfrac{\dfrac{\overset{\mathcal{D}}{All\ X\ are\ Y} \quad [J\ is\ a\ X]_i}{J\ is\ a\ Y}\ (All\ E)}{All\ X\ are\ Y}\ (All\ I)_i$$

$Some\ X\ are\ Y$ **– local soundness:**

$$\dfrac{\dfrac{\overset{\mathcal{D}_1}{J\ is\ an\ X} \quad \overset{\mathcal{D}_2}{J\ is\ a\ Y}}{Some\ X\ are\ Y}\ (Some-I) \qquad \begin{matrix}[J\ is\ an\ X]_i & [J\ is\ a\ Y]_i \\ & \overset{\mathcal{D}_3}{S}\end{matrix}}{S}\ (Some-E)_i \quad \rightsquigarrow_R \quad \begin{matrix}\overset{\mathcal{D}_1}{J\ is\ an\ X} & \overset{\mathcal{D}_2}{J\ is\ a\ Y} \\ & \overset{\mathcal{D}_3}{S}\end{matrix}$$

$Some\ X\ are\ Y$ **– local completeness:**

$$\overset{\mathcal{D}}{Some\ X\ are\ Y} \quad \rightsquigarrow_E \quad \dfrac{\overset{\mathcal{D}}{Some\ X\ are\ Y} \qquad \dfrac{[J\ is\ an\ X]_i \quad [J\ is\ a\ Y]_i}{Some\ X\ are\ Y}\ (Some-I)}{Some\ X\ are\ Y}\ (Some-E_i)$$

### 3.4. Decidability of Provability

Strictly speaking, $ND-syl^+$ does not enjoy the sub-formula property, simply because propositions in $SYL^+$ (and generally in $SYL$) do not have sub-formulas. However, both of $X, Y$ are sub-formulas of $X \multimap Y = \Pi(All\ X\ are\ Y)$, and $\vdash_{i-lint}$ does enjoy the sub-formula property.

Thus, a straightforward way to decide $\Gamma \vdash_{ND-syl^+} \phi$ is to decide $\Pi(\Gamma) \vdash_{i-lint} \Pi(\phi)$, using the known algorithm based on the sub-formula property of $\vdash_{i-lint}$. Obviously, a *direct* decision algorithm can be obtained too.

## 4. Conclusions

Clearly, the calculus presented here, in its preliminary for, constitutes only a modest first step toward the goal of $PTS$ for $NL$. The real challenge, even for this small fragment, is the incorporation into a grammar, devising a *lexicalized* $PTS$. This is currently under investigation.

## 5. Acknowledgements

We thank Larry Moss for various interactions on the paper.

## Bibliography

Moss, L.: 2005, Natural language, natural logic, natural deduction, Manuscript, in preparation

Negri, S.: 2002, A normalizing system of natural deduction for intuitionistic linear logic, *Arch. Math. Logic* 41, 789–810

Pfenning, F. and Davies, R.: 2001, A judgmental reconstruction of modal logic, *Mathematical Structures in Computer Science* 11, 511–540

Prawitz, D.: 1965, *Natural Deduction: Proof-Theoretical Study*, Almqvist and Wicksell, Stockholm

Prawitz, D.: 1971, Ideas and results in proof theory, in J. Fenstad (ed.), *Proc. 2nd Scandinavian Symposium*, North-Holland

Schroeder-Heister, P.: 2005, Validity concepts in proof-theoretic semantics, in R. Kale and P. Schroeder-Heister (eds.), *Proof-Theoretic Semantics*, Special issue of *Synthese*, to appear

# Language and Learning

# LOCALITY AND THE ORDER OF ACQUISITION STEPS

JACQUELINE VAN KAMPEN

UiL OTS
Utrecht University
Jacqueline.vanKampen@let.uu..nl

**Abstract**
Preferably, the properties of grammar can be derived from the following factors:
(i) The primary linguistic data as they are offered to the child.
(ii) A language acquisition procedure.
Hopefully, the language acquisition procedure will be compatible with plausible assumptions about the neural abilities of human beings, but that is of no immediate concern. The interaction of the primary data and the acquisition procedure can be studied by a closer look at the order of the child's acquisition steps. What does she acquire first and why? What does she acquire later and why? My main point will be that this is empirically a promising and by no means trivial approach. At the same time, I will argue against an assumption that is quite common in computational studies and also in mere grammatical studies of child language. People from Gold (1967) to Yang (2002) assume that the acquisition procedure has simultaneous access to all data at once. My point will rather be that the acquisition procedure implies a natural selection of data. The data selection procedure must predict the actual order of the acquisition steps in the various languages.

## 1. Input reduction

The procedure for first language acquisition is not confronted with all grammatical option and problems at once. The child applies a radical reduction to the mother's input. The common sense background of that reduction can be formulated as in (1).

(1) *Reduction of input to intake*
   a. Leave out what you cannot fit in.
   b. Try minimal solutions for the combinatorial restrictions in the residue.

Suppose the child has reached a point at which she is able to recognize a set of separate words with denotational content {*gone, up, car, daddy, eat*} or words with an immediate pragmatic meaning {*that, wanna, no*}. The reduction procedure in (1) will then throw out all grammatical markings: articles, copulas, auxiliaries, verbal inflections, connectives. Hence, sentences like *ain't the bear nice*; *the bear is nice, isn't he?*; *I want the bear to be nice*, are all turned into [*bear nice*]. A set of binary word constructions is the result. This at least happens in the child's actual output.

Eventually, the learning procedure will identify grammatical markings between binary combinations one at a time. This stepwise learning is an important characteristic of language acquisition. So, a proposal for a language acquisition

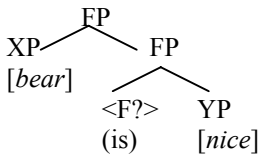
procedure should predict the reductions as they apply to an adult grammar and it should predict the linear order of acquisition steps that will follow from small, reduced utterances frames. Suppose the acquisition procedure starts with the reduction operation in (2), where <+F?> is an unidentified functional feature.

(2) *Input Reduction*
   a. *input*: substitute <+F?> for each grammatical marking still unknown.
   b. *reduction*: throw out all input sentences with more than one <+F?>.
   c. *output*: attach $F_i$ to the selectionally dominant element to the left or right

The result satisfies the *Single Value Constraint* (Clark 1992) when it happens to be that the residue from (2)b boils down to a single set defined by $F_i$. The intake to the acquisition procedure is then such that one grammatical category is singled out, identified and subsequently acquired. I hope that such a property will define natural language. Let me assume it and call the convenient outcome an evidence frame, as defined in (3) and exemplified in (4).

(3) *Evidence frame*
   a. pragmatically: an intuitively understood utterance
   b. syntactically: a binary phrase structure [ XP [ $F_i$ YP ]$_{FP}$ ]$_{FP}$
   c. semantically: fully interpretable but for a single <F?>

(4)




Functional categories are identified due to their frequency in the input and due to the fact that they cannot be and are not understood beyond grammar, i.e. beyond the grammatical relation between the phrases XP and YP.

The acquisition of the category <+F?> → $F_i$ changes the initial state and the data reduction procedure in (2) will reapply. The next grammatical category $F_{i+1}$ is singled out, etc.. The evidence frames do not follow from mere frequency of $F_i$ in the raw data set. The $F_i$ must also define a minimal frame that is fully interpretable.

The restrictive evidence frames follow from the input reduction. They offer the bootstraps for subsequent acquisition steps. It is useful, though, to realize that the evidence frames remain present and active in the adult grammar. They continue to function as the local checking domains for elementary grammatical properties. For that reason, the adult knows more or less how to schematize, c.q. creolize, his grammar. The acquisitional perspective on syntactic locality is given in (5).

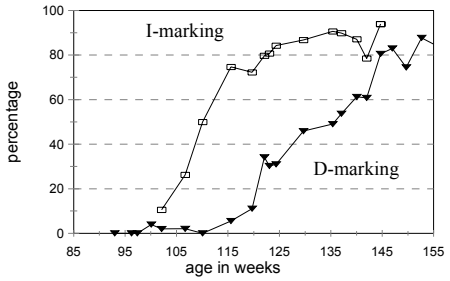
(5) a. all evidence frames in acquisition are local and binary.
   b. all grammar is acquired due to such domains.

I will present two major acquisition steps that support the reductions in (2) and (3).

## 2. Order of acquisition steps

Language acquisition overcomes the radical underspecifications that result from the initial data reduction. It proceeds by adding grammatical features within a local binary frame. The order of the acquisition steps can be shown by longitudinal graphs, see (6).

(6) Dutch Sarah: acquisition of $I^o$-marking and $D^o$-marking




The graphs in (6) represent the acquisition of finite verbs ($I^o$-marking) and determiners ($D^o$-marking) by the Dutch child Sarah (Van Kampen 2004). The graph for $I^o$-marking shows the growing percentage of grammatical predicate marking, {copula/ auxiliary/modal/finite morphology}. The graph for $D^o$-marking shows the growing percentage of grammatical argument marking, {article/demonstrative/ possessor/quantifier}.

Now the order of acquisition steps shows that Dutch Sarah applies systematic $I^o$-marking almost half a year earlier than systematic $D^o$-marking. The same order of appearance was found for English, French, and Rumanian. The amount of determiners outweighs the amount of finite verbs in the input data. Yet, children in various languages start to analyze predicate-argument structure by $I^o$-marking. The less frequent $I^o$-marking precedes the more frequent $D^o$-marking in acquisition. The order $I^o \rightarrow D^o$ must be explained. I will show how the acquisition procedure follows the *Single Value Constraint* on evidence frames as proposed in (2) when initially, sentences with both a $D^o$-marked noun and an $I^o$-marked verb are thrown out of the observation space. The feasibility of a mechanical reduction procedure was partly demonstrated by a computer simulation (Obdeijn 2004). The simulation derived child language from a child-directed input and derived an order of intake frames.

The systematic $I^o$-marking and $D^o$-marking themselves give entrance to a whole series of further acquisition steps, beginning with a grammatical decision procedure on the category membership V versus N (Van Kampen 2005). This option, chosen here for language acquisition, was implemented earlier in computational approaches to category assignment (Buszowski 1987). A general property of 'decoding' emerges as well. The successive evidence frames narrow

down to a far more precise context and the speed of acquisition increases by an order of magnitude. The subject of the $I^o$-marked predicate (finite verb) initially lacks φ-features of person/number. In a subsequent step, the φ-feature content in $D^o$, {± person, ± number} on the subject, is figured out. However, the finite verb still doesn't show the correct agreement with the subject, see (7).

(7) de clowntjes *heb* oogjes (the clowns has eyes) Sarah. week 129

One step later, the initial $I^o$-marked predicate constitutes the local evidence frame for Agreement features, the copying of the φ-features on $I^o$. The finite verb now starts showing the correct agreement. Late acquisition of agreement has been reported for various languages.

The dense succession of the acquisition steps shows that the later steps are a matter of weeks whereas the earlier steps were a matter of months, see (8).

(8)

| step $I^o$ | step $D^o$ | step $D^o$ (φ) | step $I^o$ (φ) |
|---|---|---|---|
| 20 wks | 25 wks | 5 wks | 5 wks |

The more effective acquisition relates plausibly to the more precise frame that can be used to select the input. The selection of some binary combination of content signs is far more undetermined than the distributional relation between explicit grammatical markings such as φ-features and Agreement. The later set of acquisitions is supported by a lexicon with categorial marking <+I> or <+D>. I propose that after step 1 and step 2, the EPP (subject-finite verb configuration) operates as an evidence frame.

## 3. A discovery procedure

Generative learnability theories in the 1980th were theoretical and somewhat defensive. They qualified the mathematical deduction in Gold (1967) that context-free rewriting grammars could not be identified or learned without negative data. As Wexler & Culicover (1980) argue, context free generative grammars and some transformational grammars are learnable from positive data as long as the relevant relations are sufficiently local. The main point was to argue learnability in principle for certain types of generative grammar. There was no reference to child language. The ongoing simplification of grammatical principles, pushed by Categorial Grammar, HPSG and the Minimalist Program, may re-inspire interest in their learnability. I mention four attempts into that direction. Fodor (1998), Yang (2002), Culicover & Nowak (2003) and my work with Arnold Evers (Evers & Van Kampen 2001).

Fodor (1998) and Yang (2002) assume that the child is confronted with the full variety of constructions in his language. The child meets this challenge with brilliant creativity. She comes up with all possible grammatical structures that the general theory of grammar would allow. The child's productivity in designing

possible solutions is maybe comparable with his creativity in grasping visual or musical structures or maybe with the babbling phase that precedes the construction of phonological forms. Fodor as well as Yang's learner start with a variety of grammatical structures and work towards a minimal set of grammatical structures by comparing alternative solutions. Fodor's learner is sensitive to certain key-constructions (treelets) that betray the language type and Yang's learner is sensitive to rules that are too often involved in analyses that fail. The options that they compare are assumed to be a priori present from the human brain. Yang proposes an accounting system of 'penalties' for failing rules. Yang's bookkeeping of failures and Fodor's testing system could be characterized respectively as an effective *evaluation procedure* (Yang) and as an effective *decision procedure* (Fodor). Their learners start with all options offered by the theory. Subsequently, they propose computational operations that select a language-specific grammar for the input data. Both successfully simulate how the learner zeros in on the core grammar of the language.

By contrast, I propose, like Culicover & Nowak (2003), that the young learner is unaware of the grammatical alternatives that are available in the world outside. Our learning procedure could be characterized as a *discovery procedure*. My young learner must reduce its initial attention to constructions assigned to pairs of adjacent content words and so he enters a maximally reduced observation space, as formulated in (1).

A learning procedure as in (2) that adds a grammatical feature to a category moves from a less restricted superset to a more restricted subset. The learning procedure starts with underspecifications, but the associative pressure of local contexts has a healing effect. The initial underspecifications are "blocked". Blocking effects are known from the very beginning of grammatical studies (Panini, DiScullio & Williams 1987). In general, the more specified variant blocks the less specified one. Blocking in language acquisition can be traced by longitudinal graphs as we have seen. This is a contentious issue in theories of language acquisition. Some try to reconstruct child language as subset language that is extended to the correct generalizations. Others believe that child language start with maximal generalizations and narrows down by developing subcategories (Jakobson 1942; the present perspective).

## 4. Perspective

The acquisition order is due to input-control, but definitely not always due to input frequency. Functional categories are acquired later than content words, yet their token frequency is 100 to 300 times higher than the token frequency of an arbitrary content word. Although highly frequent, functional categories can be learned only in constructions that contain content words. This is because a grammatical word $F_i$ indicates a grammatical relation between two phrases [XP [F YP]$_{FP}$] It is a word that carries no meaning beyond the syntactic relation (the word *and* does not mean 'pair', the word *but* does not mean 'objection', the word *is* does not mean

'property'). The acquisition order "content words before functional categories" is imposed by the nature of the system the child is confronted with. The same holds for $I^o$/predicate marking and $D^o$/reference marking. The $I^o$-markings are more diverse in form and less frequent in the raw input than the $D^o$-markings. Yet, $I^o$-marking precedes $D^o$-marking in acquisition.

The factual order of acquisition steps has to be established for various grammatical properties. It has to be considered whether and how that order fits the present conjecture about the hierarchy of evidence frames. For instance, the evidence frames are also effective for the subsequent learnability of scopal phenomena, like wh-marking and negation. My conjecture is that island constraints and scopal domains can be derived from the locality of the evidence frames and their dependence on the crucial terminal elements <F?> (inclusiveness). The decision proposal by Fodor and the evaluation procedure by Yang assume exclusive a priori structures, as well as procedures to compare solutions. The discovery procedure must assume that the natural input allows a reduction to local frames and a terminal string that remains informative enough in spite of the reduction. Locality and local inclusiveness of grammatical information are present to guarantee a certain type of learnability.

## Bibliography

Buszkowski, W.: 1987, Discovery procedures for categorial grammars, in E. Klein & J. van Benthem (eds.), *Categories, Polymorphism and Unification*, UvA.

Clark, R.: 1992, 'The selection of syntactic knowledge' *Language Acquisition* 2(2), pp. 83-149.

Culicover, P. & Nowak, A.: 2003, *Dynamical Grammar: Minimalism, Acquisition and Change*, University Press, Oxford.

DiScullio, A.-M. & Williams, E.: 1987, *On the Definition of Word*, MIT Press, Cambridge MA.

Evers, A. & Kampen, J. van: 2001, *E-language, I-language and the Order of Parameter Setting*, Uil OTS Working Papers 00105-S-S, Utrecht.

Fodor, J.D.: 1998, Unambiguous triggers, *Linguistic Inquiry* 29: pp. 1-36.

Gold, E.M.: 1967, Language identification in the limit, *Information and Control* 10, pp. 447-474.

Jakobson, R.: 1942, *Kindersprache, Aphasie und Allgemeine Lautgesetze*. Uppsula.

Kampen, J. van: 2005, Language specific bootstraps for UG categories, *International Journal of Bilingualism* 9-2, pp. 253-277.

Neeleman, A. & Koot, H. van de: 2002, The configurational matrix, *Linguistic Inquiry* 33-4, pp. 529-574.

Obdeijn, A.: 2004, *Taalverwerving door Kinderen en Machines*, Master thesis University of Amsterdam, ILLC.

Yang, C.D.: 2002, *Knowledge and Learning in Natural Language*, U.P: Oxford.

# Semantic Universals

# PROHIBITIVES: WHY TWO THIRDS OF THE WORLD'S LANGUAGES ARE UNLIKE DUTCH[1]

JOHAN VAN DER AUWERA

Center for Grammar, Cognition and Typology

University of Antwerp
johan.vanderauwera@ua.ac.be

**Abstract.** Most languages of the world have prohibitive constructions that use a negative marker than is more or less dedicated to this construction, rather than a negative that also serves in declarative constructions. The explanation has to appeal to aspect. Declarative negation is inherently stative, but prohibition is inherently dynamic. The negative of the declarative is therefore not well suited to appear in prohibitives.

This is an exercise in "semantic typology": it investigates a semantically interesting property of human language as such (prohibition) through the window of an analyis of a sample of the world's languages.

## 1. Introduction

A prohibitive construction is a negative imperative construction, i.e., an imperative construction that appeals to the hearer(s) to establish or to maintain a negative state of affairs. It has been claimed that prohibitive constructions often do not use the negative marker found in negative declaratives. In section 2, I show that this claim is correct. In section 3, I offer a semantic explanation.

## 2. Languages prefer 'prohibitive markers'

### 2.1. Prohibitive constructions and prohibitive markers

Let 'prohibitive' be the one word term for 'negative imperative'. I will speak about both 'prohibitive constructions' and 'prohibitive markers'. Let me explain the former notion first. A prohibitive construction is that construction, whatever its structure, that is conventionally used to express a prohibition. It can be illustrated with Dutch (1b).

[1] This paper is based on van der Auwera (in print).

| (1) | a. | Beweeg! | b. | Beweeg | niet! |
|---|---|---|---|---|---|
| | | move.IMP.2 | | move.IMP.2 | NEG |
| | | 'Move!' | | 'Don't move!' | |

| | c. | Hij | beweegt | niet. |
|---|---|---|---|---|
| | | he | move.IND.PRS.3SG | NEG |
| | | 'He doesn't move!' | | |

In Dutch, the prohibitive construction is straightforward: it employs the same verb form as the one found in (positive) imperatives, and the negation (*niet)* is the one found used in declaratives (illustrated in 1c).

Mandarin is different in this respect. First of all, in declaratives Mandarin has a few sentential strategies, the three most common ones employing the markers *bù, méi* or *méiyou* (examples from Li and Thompson 1981: 417-8).

| (2) | a.. | 他 | 不 | 念 | 书. |
|---|---|---|---|---|---|
| | | Tā | bu | niàn | shū. |
| | | 3SG | NEG | study | book |
| | | '(S)he does not study.' | | | |

| | b. | 他 | 没 有 | 开 | 门. |
|---|---|---|---|---|---|
| | | Tā | (méi)yǒu | kāi | mén. |
| | | 3SG | NEG | open | door |
| | | '(S)he didn't open the door.' | | | |

*méi*(*yǒu*) does not surface in prohibitive constructions, and neither does *bu*, at least not by itself. Instead we mostly find the dedicated markers *buyào* and *bié.* They are grammatical markers with just that function (*bié*) or with that function as one of its main functions (*buyào*) (examples from Li and Thompson 1981: 455, Yip and Rimmington 1977: 88).

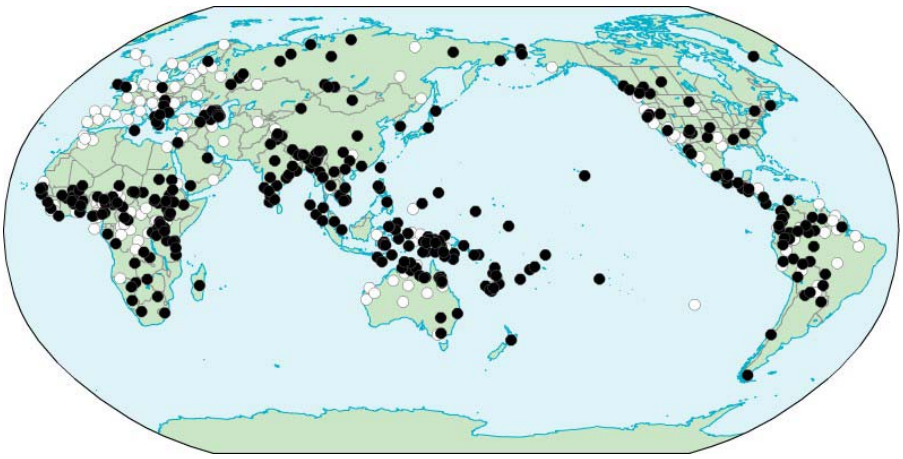
| (3) | a. | 动! | b. | 别 / 不 要 | 动! |
|---|---|---|---|---|---|
| | | Dòng! | | Bié / Buyào | dòng! |
| | | move | | PROH/PROH | move |
| | | 'Move!' | | 'Don't move!' | |

The (degree of) dedication of the negative markers *bié* and *buyào* to the prohibitive construction will be honored with a technical term: I will call them 'prohibitive markers'.

### 2.1. A universal preference for prohibitive markers

It turns out that the Mandarin situation is more typical for the world's languages than the Dutch one. This has been pointed out before, since at least Schmerling

(1982: 202). The largest and most representative sample on which such claims have been based are those of van der Auwera and Lejeune (2005b). They looked at second singular prohibitive constructions of 495 languages. They concluded that 327 of those 495 languages have a more or less dedicated prohibitive marker as their only or main strategy. That we are dealing with a truly world wide phenomenon is suggested by Map 1 below, based on the map in van der Auwera and Lejeune (2005b: 292-3). There is at least one exceptional area, viz. Western Europe—an areal claim that has also been hinted at in the literature (since at least Sadock and Zwicky 1985: 175-177).



Map 1. Prohibitive markers in second person singular prohibitive constructions
● Prohibitive marker as unique or main strategy
○ No prohibitive marker as unique or main strategy

## 3. Why do languages prefer 'prohibitive markers'?

Explanations can be formal or functional. In generative quarters most work (by and inspired by Zanuttini 1997) does not directly address the preference for prohibitive markers, but rather the fact that in some languages the declarative negative marker does not combine with the imperative. This is due to the initial focus on Romance languages, in which the negative is unremarkable but the verb often has to be subjunctive (or infinitival).

(4) Spanish

| a. | Canta! | b. | No | cantes! |
|---|---|---|---|---|
| | IMP.2Sg | | NEG | sing.SUBJ.PRES.2SG |
| | 'Sing!' | | 'Don't sing!' | |

Of course, the generative work is relevant for my question too, for part of the

reason for the universal preference for a prohibitive marker might be the dislike of the combination of the declarative negative and the imperative.

Formalist explanations have tended to derive the unavailability of the combination of the declarative negative and the imperative from other properties of the declarative negative and the imperative. Relevant features have included (i) the position of the negative relative to the verb , (ii) the question whether or not the negative is a clitic, and (iii) the question whether or not the verb form used for commands is a dedicated imperative or whether it instantiates just one use of a multifunctional category, which is then said to supply "suppletive" imperatives. Though these factors may well be relevant for specific languages, as an explanation for the universal preference for prohibitives, they must be discounted. It can easily be shown that prohibitive markers appear independently of any formal and positional properties. There is also no dependence on whether or not the imperative is dedicated. This was a claim by Dooley Collberg and Håkansson (1999: 32), arrived at on the basis of their 15 language sample. I can replicate it on the basis of the descriptions in van der Auwera and Lejeune (2005a, 2005b). For 473 languages we had data on the dedicatedness of the second singular imperative and on the existence of a second singular prohibitive marker. Of these the majority have a dedicated imperative. But both in the subset with dedicated imperatives and the one without, prohibitive markers are preferred, and even in roughly the same proportion (see Table 1).

| | | Prohibitive marker | |
|---|---|---|---|
| | | + | - |
| Morphologically | + | 236 | 131 |
| dedicated IMP.2SG | - | 78 | 28 |

*Table 1.* The second singular morphologically dedicated imperative and the second singular prohibitive marker

What is it then that makes prohibitive constructions prefer prohibitive markers? I propose that we are dealing with an illocutionary garden path effect. For this I have to assume that the declarative negatives are the most frequent ones, i.e., more frequent than either imperative negatives or interrogative negatives, and that it is important for languages to make clear whether the direction of fit—to use a speech act term—is word to world (declaratives) or world to word (imperatives). If these assumptions are correct, then there is a certain risk that a occurrence of the negative marker that occurs in declaratives will initially be taken as a sign of a declarativeness, even when it turns out that the speech act is not declarative. Of course, the hedge of "a certain risk" is important, for quite some languages do take that risk.

To clarify this further, we can bring in considerations of aspect. As many linguists have remarked, also in typology (most recently Miestamo 2003: 185), it is a property of negative declaratives that they are inherently stative.

Positive declaratives, on the other hand, may be stative or dynamic. Consider the positive declaratives in (6a,c,e).

| | | | | |
|---|---|---|---|---|
| (6) | a. | John was at home. | b. | John wasnt at home. |
| | c. | It rained. | d. | It didn't rain. |
| | e. | John ran away. | f. | John didn't run away. |

(6a) is stative and (6c) and (6e) are dynamic: (6c) is a process and (6e) is an action. But consider now their negations in (6b,d,f). All of these are stative: (6b) is of course just as stative as its positive counterpart. But (6d) and (6f) are stative now, the simple reason being that nothing happened, the process didn't materialize and neither did the action. I can therefore conclude that the most frequent use of the negative, that of the declarative negative, is stative. Or again, the most frequent use of the negative can be paraphrased with 'it is not the case that'. The negative imperative or prohibitive, however, is not stative. On the contrary, it is an appeal for action, either of discontinuing what is going on or of taking care that some new state of affairs does not materialize. This use does not support any 'it is not the case that' paraphrase at all. The appropriate paraphrase is rather 'let it not be the case that'. The negative of the prohibitive is thus crucially different from the most frequent use of the negative. There is thus—and I use a hedge again—a "certain need" for reflecting this difference in a direct way, most clearly so with the help of a more or less dedicated prohibitive marker.

What happens then in the languages that do use the same negative as the one they use in declaratives? I see at least three scenarios. First, the marker abstracts from the speech act type of the utterance. This is the case of Dutch *niet*. It is freely used in any type of speech act. Second, the marker is the declarative negative simply because the prohibitive construction is or was declarative too. Consider the Oceanic language Mussau-Emeria.

(7) Mussau-Emeria (Ross 2002: 165)

| | | | | | | |
|---|---|---|---|---|---|---|
| Karika | u | mene | nama | asi | eteba | o. |
| NEG | 2SG | again | eat | taro | SG | that |

'You will not eat that taro!'
'Don't eat that taro!'

Not only is the *karika* gram the one that we find in a declarative, the whole sentence allows a declarative reading, more particularly, a future one, meaning 'You will not eat that taro!'. This is actually a common situation and one could either say that the language in question does not really have a prohibitive construction but uses the negative future instead or that the pattern is vague between a declarative future and a prohibitive reading.

The third scenario is that of the Spanish subjunctive (or infinitive). In this case, the strategy is a conventionalization of the description of the content of the desired state of affairs. What one wishes when prohibiting the hearer not to sing is that the speaker would not sing—a subjunctive also in English— i.e.,

an irrealis state of affairs of no singing. In the case of the conventionalization of the subjunctive turn of Spanish, it must have taken place a long time ago, for this strategy has been around since the earliest documents of Latin. But it would not suffice to explain the Spanish subjunctive by merely saying that it is a relic from Latin. One must conclude that framing the prohibition with the indirect strategy of describing only the content of the prohibition must have the independently commendable effect of softening the prohibition or, to vary on Horn (1991: 97), of "cushion[ing] the iron fist" of prohibition "in the velvet glove" of the description of what is merely wished for.

## Bibliography

Dooley Collberg, S. and Håkansson G.: 1999, Prohibition: negative imperatives and the parametric typology of negation, *Lund University, Department of Linguistics, Working Papers* 47: 25-37.

Horn, L.R.: 1991, Duplex negation affirmat: the economy of double negation, in L.M. Dobrin et al (eds.), *Papers from the 27th regional meeting of the Chicago Linguistic Society. Part two: the parasession on negation*, pp 180-96, Chicago Linguistic Society, Chicago

Li, C.N. and Thompson, S.A.: 1981, *Mandarin Chinese*. *A functional reference grammar*, University of California Press, Berkeley

Miestamo, M.: 2003, *Clausal negation. A typological study*, Ph.D. thesis, University of Helsinki

Ross, M.: 2002a, Mussau, in J. Lynch, et al (eds.), *The Oceanic languages*. pp.148-65, Curzon, Richmond

Sadock, J.M. and Zwicky, A.: 1985, Speech act distinctions in syntax, in: T. Shopen (ed), *Language typology and syntactic description*. *Volume I. Clause structure*, pp 155-196, Cambridge University Press, Cambridge

Schmerling, S.S.: 1982, How imperatives are special, and how they aren't, in R. Schneider et al (eds.), *Papers from the parasession on nondeclaratives,* pp 202-18, Chicago Linguistic Society, Chicago

van der Auwera, J.: In press, Why languages prefer prohibitives?, *Journal of Foreign Languages*

van der Auwera, J. et al: 2005a, The morphological imperative, in M. Haspelmath et al (eds.), *The World Atlas of Language Structures,* pp. 286-9, Oxford University Press, Oxford

--- 2005b, The prohibitive, In: M. Haspelmath et al (eds.), *The World Atlas of Language Structures*, pp 290-3, Oxford University Press, Oxford

Yip, P.-C. and Rimmington, D.: 1977, *Chinese: An essential grammar*. Routledge, London

Zanuttini, R.:1997, *Negation and clausal structure: a comparative study of Romance languages,* Oxford University Press, Oxford

# ASSOCIATION: A CROSS-LINGUISTIC EXPERIMENT

DAVID GIL

Department of Linguistics

Max Planck Institute for Evolutionary Anthropology
gil@eva.mpg.de

In a series of publications, I have argued that Riau Indonesian exhibits a number of syntactic and semantic features that characterize it as typologically exceptional. The question arises whether Riau Indonesian is truly exceptional, or whether its apparently exceptional properties are a mere artefact of a particular descriptive approach. In order to answer this question, it is necessary to compare Riau Indonesian to other languages through the same eyes, using the same objective and rigorous yardsticks. This paper proposes one such yardstick, in the form of a psycholinguistic experiment designed to elicit truth-value judgments in different languages.

At the heart of the semantic analysis of Riau Indonesian is the claim that when two expressions X and Y with meanings P and Q respectively are combined, the meaning of the collocation X Y is derived from that of its constituent parts by means of the association operator, A (P, Q), which says that the meaning of X Y is associated in an unspecified way with the meanings of X and Y respectively. For example, if *ayam* means 'chicken' and *makan* means 'eat', *ayam makan* means A ( CHICKEN, EAT ), or anything that has to do in some way with 'chicken' and with 'eat'. In particular, the semantic representation A ( CHICKEN, EAT ) lacks any specification of thematic roles: the chicken could assume the role of agent, patient, or whatever might make sense in the context of the utterance.

This paper presents the results of an experiment designed to measure, objectively across a variety of languages, the availability of *apparently associational interpretations*: interpretations that appear to be obtainable from the association operator without reference to thematic roles or other semantic categories. Two kinds of apparently associational interpretations are sought: (a) those in which what looks like a bare noun preceding a bare verb is interpreted as the patient (rather than the agent); and (b) those in which what looks like a bare noun in construction with a bare verb is interpreted as an oblique argument or even a non-argument (in the absence of prepositions or other such markings). The experiment presents subjects with a sentence in the target language and two pictures; subjects are asked which of the two pictures is

best described by the sentence. The experiment is ongoing; as of August 2005, over 1000 subjects in a dozen languages had been tested.

While non-isolating languages have near-zero availability of apparently associational interpretations, isolating SVO languages generally allow apparently associational interpretations to some extent, thereby setting such languages apart from most others. However, amongst themselves, isolating SVO languages exhibit substantial cross-linguistic variation with respect to the availability of apparently associational interpretations. In this regard, the position of Riau Indonesian amongst isolating SVO languages is not exceptional: it falls in the mid range of Malayic languages, and in the mid-range of other isolating languages, in fact with substantially lower availability of apparently associational intepretations than other West Malayo-Polynesian languages such as Minangkabau and Sundanese.

# CASE AND STRENGTH

HELEN DE HOOP

CLS, Linguistics
Radboud University
Nijmegen
H.deHoop@let.ru.nl

In the functional-typological literature two main functions of case-marking are distinguished. One motivation for case-marking is *disambiguation*, i.e. the need to distinguish between the arguments of a two- or three-place relation. Another widely attested function of case involves the *expression* or *identification* of specific semantic information. In this paper I will investigate the two functions of case-marking to see where they converge and diverge with respect to the semantic features of arguments that are case-marked. I will focus on the 'strength' of the arguments in relation to their case-marking.

## 1. Introduction

The 'strength' of arguments can be viewed as a function of their 'discourse prominence' or of their degree of 'typicality' as a full-fledged argument. Under both perspectives, it can be argued that semantic features such as animacy and definiteness contribute to the strength of grammatical arguments. The aim of this paper is to explore the relation between the strength of nominal arguments and the 'meaning' of case.

## 2. A functional perspective on case marking

I assume that in ergative-absolutive systems ergative case is assigned to the first argument *x* of a two-place relation *R(x,y)*, while in nominative-accusative systems accusative case is assigned to the second argument *y* of a two-place relation *R(x,y)*. Morphologically unmarked (abstract) case is analysed as the absence of case. Nominative case in nominative-accusative systems, and absolutive case in ergative-absolutive patterns, is often unmarked. In many languages, ergative and accusative case are assigned only (or mainly) in transitive sentences, leaving the subject of an intransitive sentence without case. This can be functionally explained (Comrie 1989, Aissen 2003). In order to differentiate the subject from the object it is not necessary to mark them both; a case marker on one of them already serves to distinguish the two arguments. So, when *x* and *y* are the two arguments of a transitive clause, and *z* is the one argument of an intransitive clause, then the picture that emerges under this function of case-marking is as follows. In ergative-absolutive languages *x* gets ergative case, while *y* stays unmarked (called

'absolutive' case, which is often the absence of case). In nominative-accusative case systems, *y* is marked, while *x* stays unmarked (called 'nominative' case). The only argument *z* of intransitive clauses stays unmarked in both case systems (thus, it patterns with the transitive object in ergative-absolutive languages and with the transitive subject in nominative-accusative languages). This function of case-marking is generally referred to as the *distinguishing* or *discriminating* function.

The distinguishing function of case is not its only function, however (Song 2001). Morphological cases are generally considered to express some kind of specific (e.g., thematic) information of the noun phrase that bears the case morphology. For instance, dependent on its case, a noun phrase can refer to the goal, the agent, or the experiencer of an event, or it is interpreted with respect to a certain location or direction in space. This function of case marking is usually referred to as the *identifying* or *indexing* function of case. Both functions of case-marking seem to manifest themselves within the different (more syntactic and more semantic) domains of case-marking.

## 3. Case and argument strength

De Hoop and Narasimhan (2005) point out that, dependent on which function of case-marking is dominant in a certain language, different arguments may get case-marked, the 'strong' or the 'weak' ones. Cross-linguistically, the strength of the arguments seems to influence case-marking. In de Hoop (1996) I suggested that in languages with differential case-marking, subjects and objects that are 'strong' are likely to be overtly case-marked. However, this does not always hold. In fact, sometimes the 'weak' rather than the strong arguments receive overt case-marking (cf. Aissen 1999, De Swart 2003, De Hoop and Narasimhan 2005).

The question is how we measure the 'strength' of arguments, since languages may vary in this respect. One notion that seems relevant in this respect is 'discourse prominence'. Legendre *et al.* (1993) use Optimality Theoretic constraints such as "High-prominence arguments receive $C_1$" and "Low-prominence arguments are not case-marked $C_1$ and $C_2$", where $C_1$ in their framework refers to both nominative and ergative, and $C_2$ to both accusative and absolutive. According to Legendre *et al*., the one argument of an intransitive clause is always high-prominent, and the two arguments of a transitive clause too. They write high-prominent arguments with capital letters. Thus, the subject and object of a transitive clause can be written as *X* and *Y*, and they are universally marked $X_1Y_2$ (where the subscript indicates the type of case). Legendre *et al.* (1993) argue that passivization applies when the input transitive clause has a low-prominent subject (*xY*) , while antipassives are the result of an input with a low-prominent object (*Xy*). However, they do not account for the fact that passives are found more often in nominative-accusative languages, while antipassives are found more often in ergative languages (Malchukov, to appear).

Moreover, if we compare two transitive clauses, e.g., *Jane hit Jacky* on the one hand, and *Jane was drinking wine* on the other, then intuitively *wine* is less prominent in the discourse than *Jacky*. Hence, we should use the input *Xy* for the sentence *Jane was drinking wine* although syntactically this is not an antipassive

construction in English. So, another way of measuring the 'strength' of an argument is by looking at its typicality as a full-fledged argument (of a transitive clause) (Hopper and Thompson 1980). Animate and specific arguments are more often realized as 'real' arguments than inanimate and non-specific arguments. Note that in fact the different ways of measuring the 'strength' of arguments point in the same direction: arguments which are high-prominent in the discourse are usually animate and specific, whereas arguments that are 'typical' arguments are usually animate and specific as well.

## 4. Markedness

One of the fundamental insights of functional typology is that the most typical instances of a certain category are the least likely to be marked (cf. Silverstein 1976). For example, following Comrie (1989), Aissen (2003) notes that an object without case morphology is used for a typical (i.e., semantically unmarked) object in languages that show Differential Object Marking (DOM). A semantically unmarked object is a 'weak' object, for example a non-specific object. If, on the other hand, the object is specific, the meaning is considered *marked* (for an object), hence the object will be case-marked. Hindi provides an example of this type of case-alternation.

| | | | | | |
|---|---|---|---|---|---|
| (1) | us=ne | ek | bakraa | becaa | |
| | he=ERG | one | goat | sold | |
| | "He sold a goat." | | | | |
| (2) | us=ne | ek | bakre=ko | | becaa |
| | he=ERG | one | goat=ACC | | sold |
| | "He sold the goat." | | | | |

Næss (2004), however, points out that, from a language typological point of view, one could also claim that the object in (2) is typical rather than the one in (1). Her argument is that in languages that *have* transitive constructions at all, sentences like (2) are always transitive, whereas languages differ in whether they express sentences like (1) as transitive constructions. Take for example the alternation between (3) and (4) from Greenlandic Eskimo (Bittner 1988):

| | | | |
|---|---|---|---|
| (3) | Jaaku | arna-mik | tuqut-si-v-uq |
| | Jacob | woman-INSTR | kill-AP-IND-3sNOM |
| | "Jacob killed a woman." | | |
| (4) | Jaaku-p | arnaq | tuqut-p-as |
| | Jacob-ERG | woman | kill-IND-3sERG/3sNOM |
| | "Jacob killed the woman." | | |

Note that in (3) the 'direct object' or the *y* argument is weak/non-specific, whereas it is strong/specific in (4). Yet, only (4) is a true transitive construction with ergative case on the subject and both subject and object agreement on the verb, whereas (3) is in fact an intransitive, more specifically an antipassive construction

and its only 'true' argument (the subject) is therefore unmarked for case, whereas the 'object' is marked with oblique (instrumental) case. The question is how one can maintain that a typical direct object in a transitive clause (i.e., its meaning) is non-specific, if this type of object is very often *not a grammatical object at all* from a morpho-syntactic perspective point of view (i.e., its form).

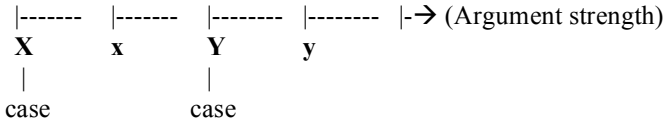
The standard views on this issue largely diverge. Whereas Comrie (1989), Aissen (2003), and De Swart (2003) claim that (1) is an example of an unmarked transitive, Hopper & Thompson (1980), Legendre *et al.* (1993), and Næss (2004) take the opposite view, and claim that sentences like (2) where both arguments are high-prominent in the discourse are truly transitive.

'Transitive' verb phrases consisting of a verb and a noun phrase can thus be composed in different ways. For instance, the verb can be straightforwardly transitive (<e,<e,t>>) in the sense that it denotes a relation between two equal arguments, or the verb is formally intransitive (or detransitivized by an antipassive marker) and its object functions more or less as a predicate modifier (type <<e,t>,<e,t>>) (cf. de Hoop 1996). Languages can differ in what counts as 'more marked'.

## 5. So, what does case mark?

Case-marking in order to *identify* arguments can be illustrated as follows:

**Figure 1:** The identifying function of case (de Hoop and Narasimhan 2005)




Clearly, in its identification function, case-marking identifies *strong* arguments as these function as the typical arguments of a transitive verb (the 'true' subjects and the 'true' objects). So, under the identifying function case is expected to mark the strong subjects and objects in a transitive clause. Note that the identification function is not limited to the core arguments of transitive sentences. Aristar (1997) points out that the strength of noun phrases (in particular, animacy and definiteness) influences other types of case-marking as well. For example, in Yidiny inanimate nouns are marked by locative case, whereas animate nouns that get a similar locative meaning, are marked by dative case (Dixon 1977):

(5) dajbu-: wunaŋ djaŋga
ground-LOC exist hole
"There are holes in the ground."

(6) buŋa:-nda wunaŋ djaŋga
woman-DAT exist hole
"There are holes in women."

Aristar argues that it is not animacy per se that plays a role, but rather discourse prominence in general. The following dialogue shows that when a locative inanimate noun phrase is high-prominent it is marked by dative case (Dixon 1977).

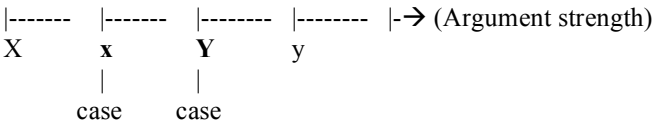
(7) ŋayu djana:-n naru walba-:
I stood top stone-LOC
"I stood on top of a stone."

(8) nundu djana:-n naru walba:-nda
you stood top stone-DAT
"Oh, you stood on top of the stone!"/ "It was a stone you stood on top of."

Clearly, the identifying function of case can be argued to mark strong noun phrases, independent of their semantic role or syntactic function.

As was pointed out above, differential case-marking can also be employed to distinguish between subjects and objects. Since subjects are usually stronger than objects, obviously, differential case-marking on the basis of distinguishability marks weak subjects rather than strong ones because the weak ones are 'closer' to the objects in strength. On the other hand, strong objects get case-marked and not the weak ones, because in the case of objects, the strong ones are 'closer' to the subjects. This is illustrated as follows:

**Figure 2:** The distinguishing function of case (de Hoop and Narasimhan 2005)




Comparing figures 1 and 2, an essential difference between the two functions of case-marking is revealed. While the identifying function explains case-marking of the strong subject and the strong object (the *X* and the *Y*), the distinguishing function explains case-marking of the weak subject and the strong object (the *x* and the *Y*).

## 5. Conclusion

When case-marking is merely used to distinguish two arguments in a transitive clause, it is sufficient to mark only one of the two arguments. By contrast, when case-marking is used to identify the strength or prominence of an argument, it may apply to each argument independently, both the subject and the object of a transitive clause, as well to other noun phrases. Here, the two functions of case-marking diverge. Moreover, while the identifying function explains case-marking of the strong subject and the strong object, the distinguishing function explains case-marking of the weak subject and the strong object.

**Acknowledgements**

I gratefully acknowledge the Netherlands Organisation for Scientific Research, NWO (grants no. 051.02.070, 220-70-003 and DN 30-609). I would like to thank Jaap van der Does, Andrej Malchukov, Bhuvana Narasimhan, and Peter de Swart, for helpful discussions.

**Bibliography**

Aissen, J.: 2003, Differential Object Marking: Iconicity vs. Economy. In *Natural Language and Linguistic Theory* 21, 435-483

Aristar, A.R.: 1997, Marking and Hierarchy, Types and the Grammaticalization of Case-Markers. In *Studies in Language* 21, 313-368.

Bittner, M.: 1988,

Comrie, B.: 1989, *Language universals and linguistic typology*. University of Chicago Press.

Dixon, R.M.W.: 1977, *A grammar of Yidiny*. Cambridge University Press.

Van Geenhoven, V.: 1996

de Hoop, H.: 1996: *Case Configuration and Noun Phrase Interpretation*. Garland.

de Hoop, H. & B. Narasimhan: 2005, Differential case-marking in Hindi. In: M. Amberber & H. de Hoop, H. (eds.). *Competition and Variation in Natural Languages: the Case for Case*. Elsevier.

Hopper, P.J. and S.A. Thompson: 1980, Transitivity in Grammar and Discourse. *Language* 56, 251-299.

Legendre, G., W. Raymond, and P. Smolensky: 1993, An Optimality-Theoretic Typology of Case and Grammatical Voice Systems. *Proceedings of the 19th Meeting of the Berkeley Linguistics Society*. Berkeley Linguistics Society.

Malchukov, A.L.: to appear, Transitivity parameters and transitivity alternations: constraining co-variation. In: L. Kulikov, A. Malchukov, and P. de Swart (eds.), *Case, valency and transitivity: a cross-linguistic perspective*. John Benjamins.

Næss,Å. : 2004, What Markedness Marks: The Markedness Problem with Direct Objects. *Lingua* 114, 1186-1212.

Silverstein, M.: 1976, Hierarchy of features of ergativity. In: R.M.W. Dixon (ed.) Grammatical categories in Australian languages. Australian Institute of Aboriginal Studies.

Song, J.J.: 2001, Linguistic Typology: Morphology and Syntax. Longman.

de Swart, P.: 2003, The Case Mirror. MA-Thesis, University of Nijmegen

# HOW MUCH LOGIC IS BUILT INTO NATURAL LANGUAGE?

ED KEENAN

Linguistics
UCLA
keenan@humnet.ucla.edu

**Query** Does knowing a natural language (English, Japanese, Swahili,...) imply knowing any logic?

**The Query is reasonable** (First Order) Predicate Logic ($PL_{=}$) is a "Universal Grammar" for the languages of Elementary Arithmetic, Euclidean Geometry, Set Theory, .... It defines their expressions, their semantic interpretations, and proofs, that syntactically characterize the boolean semantic entailment relation.

## 1. Properties of PL overtly present in Natural Language (NL)

### 1.1. Function Symbols (F1s, F2s,...) and Naming Expressions (F0s)

**PL:** $+$ and $\times$ are F2s, squaring $^2$ is an F1: 2, 3, $3^2$, $(3^2+2)$, $(3^2+2)^2$, ...

**NL:** kin terms are F1's: *the dean, the mother of the dean, the mother of the mother of the dean,*.... These are easier to understand if we vary the function expression: *the wife of the employer of the mother of the dean, etc.,* ...

**Recursion =** the values of a function lie in its domain, so its application iterates Not limited to possessive constructions. In children's rhymes and songs:

**Relative clauses** *This is the house that Jack built, This is the malt that lay in the house that Jack built, This is the rat that ate the malt that lay in the house* ...

**Prepositional phrases** *There's a hole in the bottom of the sea, There's a log in the hole in the bottom of the sea, There's a bump on the log in the hole...*

**Compositionality** meaning of a derived expression a function of those it is derived from: '$(2+3)$' denotes the value of the function denoted by '$+$' at the numbers denoted by '2','3'.

**A Fundamental Similarity** PL and NL are recursive, compositional systems. They build infinitely many non-synonymous expressions from a finite list.

**Leading Question of Md Linguistics:** Account for how we produce and understand arbitrarily many novel expressions in NL. Recursion + Compositionality a partial answer

Recursion (self application) is a "statistical accident." Most functions don't iterate: *The height of the dean, #the height of the height of the dean,* ...

### 1.2. Predicate-Argument Formulas (FMs) /Sentences (Ss)

**PL** Simple FMs = Predicate + Names. '$2 > 1$', $2^2 = (3 + 1)$.

**NL** abundant: P1s ≈ *sleeps*; P2s ≈ *praises*; P3s ≈ *gives*;

Arguments are often **asymmetrically** related: In PL $2 > 1$ and $1 > 2$ both make sense (but differ in truth value). *I wrote that poem* is natural, *That poem wrote me* is nonsense. The first argument of *write* is its **Agent**, the second its **Patient**.

The second argument of a P2 may be referentially bound to the first, but not conversely:

Ben washed/punished himself *Himself (Heself) washed/punished Ben

P2s in NL may fail to be isomorphic. *Ben washed the car* passivizes to *The car was washed by Ben*. But *Ben has a car* does not passivize: **The car is had by Ben*.

### 1.3. Boolean operations

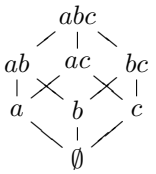
**In PL** *and*, *or*, and *not* build FMs from FMs; they denote boolean functions: *and* is a binary greatest lower bound (glb) operator, noted $x \wedge y$; *or* a binary least upper bound (lub) operator, noted $x \vee y$, *not* is a complement operator, noted $\neg x$. Writing $TV(\phi)$ for the truth value of $\phi$, we have $TV(\phi \& \psi) = TV(\phi) \wedge TV(\psi)$, $TV(\phi \text{ or } \psi) = TV(\phi) \vee TV(\psi)$ and $TV(\text{not } \phi) = \neg(TV(\phi))$.

**Negation** Present in all NLs (Dryer 2005). *Bill* isn't *a linguist.*

**'and', 'or' and 'neither ... nor ...'** combine with expressions in most categories (in PL they only connect FMs); *Both John and Bill either laguhed or cried.*

**Boolean Lattices** (B,$\leq$) are distributive, complemented lattices. $\leq$ is a boolean partial order:

BL8




BL2


Here $x \leq y$ iff $x = y$ or you can move up along edges from $x$ to $y$. $x \wedge y$ is the "highest" element which is $\leq x$ and $\leq y$. $x \vee y$ is the lowest element that both $x$ and $y$ are $\leq$ to. BL2 is the boolean lattice of truth values, which FMs denote.

**Gen 1** *The set in which expressions of a category C denote is a boolean lattice (B, ≤), supporting that the boolean operations are "properties of mind" (Boole 1854).*

**Gen 2** *Modifiers are usually restricting:* tall student ≤ student*, that is, all tall students are students, all skillful doctors are doctors, etc.*

**Variable usage in NL:** *and* may = *and then*, as in *Flo got married and got pregnant ≠ Flo got pregnant and got married*, or *and as a result*, as in *John drank too much and got sick*. But not always: *Flo is 6 feet tall and studies biology = Flo studies biology and is 6 feet tall*. Usage in logic abstracts from this variation to yield $(P \& Q)$

is true iff $P$ is and $Q$ is, whence the semantic symmetry: $P\&Q = Q\&P$. Similarly with *or*, which is sometimes intended as exclusive, as in *John either laughed or cried* (? but not both). But not always: *Do you have two nickels or a dime?* must be answered 'Yes' if you have both two nickels and a dime.

**Quantification in PL** $Qx\phi$ is a FM, where $\phi$ a FM, x a variable and $Q$ is either the universal quantifier *all*, noted $\forall$, or the existential quantifier *there exists*, noted $\exists$.

$\forall x(x^2 > x)$ 'The square of any number is greater than that number'

$\exists x(Even(x)\& Prime(x))$ 'There is a number x which is both even and prime'

Semantically $\forall$ an arbitrary glb operator, and $\exists$ an arbitrary lub operator. E.g. $TV(\forall x(x^2 \geq x)) = TV(0^2 \geq 0) \wedge TV(1^2 \geq 1) \wedge (2^2 \geq 2) \wedge \ldots$ Writing $TV(\phi[^x/_b])$ for the truth value of $\phi$ when the variable $x$ is set to denote $b$, we see that $\forall$ = "AND writ big"; $\exists$ = "OR writ big." $TV(\forall x\phi) = \bigwedge\{TV(\phi[^x/_b]) \mid b \in E\}$ and $TV(\exists x\phi) = \bigvee\{TV(\phi[^x/_b]) \mid b \in E\}$

PL ties variable binding to quantification. It is enlightening to separate them, as in $ALL(x.\phi)$ where $(x.\phi)$ is a P1 built from a FM (P0) by prefixing the variable $x$. Then Qs combine directly with P1s to form FMs (P0s). (Read $(x.\phi)$ as $\lambda x.\phi)$.

**NL Universal Quantification** Present in all NLs (knowledgeable conjecture, kc; Stassen 2005, Gil 2005). *All* cats are black; The students have *all* left;

**Existential Quantification** All NLs may assert and deny existence (kc): *There are / (aren't any) children in the park.*

## 2. Logical Properties covertly present in NL (Whorf-Sapir Hypothesis)

Knowing English implies knowing the distribution of NPI's (negative polarity items)—e.g., *ever* and *any*, whose presence is licensed by overt negation, as in (1), but also by certain NPs in subject position, as in (2):

(1) a. John hasn't ever been to Pinsk  a'. John didn't see any birds on the walk
b.*John has ever been to Pinsk  b'.*John saw any birds on the walk.

(2) a. No student here has ever been to Pinsk
a' Neither John nor Mary knew any Russian
b. *Some student here has ever been to Pinsk
b' *Either John or Mary knew any Russian.
c. Fewer than five / *More than five students here have ever been to Pinsk
d. At most / *At least two students here have ever been to Pinsk

**query** Which NPs license NPI's? What do they have in common with negation?

**Gen 3** *NPI licensers are expressions which denote monotone* ***decreasing*** *functions*

**Def** Let (A,$\leq$) and (B,$\leq$) be posets, $F$ a function from A into B. Then

a. $F$ is *increasing* iff for all $x, y \in$ A, if $x \leq y$ then $F(x) \leq F(y)$.

b. $F$ is *decreasing* iff for all $x, y \in$ A, if $x \leq y$ then $F(y) \leq F(x)$.

**Test for Increasingness (↑):** if all Ps are Qs and X is a P, therefore X is a Q. Ex: 'some poet' is ↑: Suppose all Londoners drink stout and some poet is a Londoner. Therefore some poet drinks stout.

**Gen 4** *Virtually all syntactically underived NPs are ↑: Proper Names (*Ned*,* Gail*), pronouns (*he*,* she*,* they*), demonstratives (*this*,* those*).*

**Gen 5** *The closure of Proper Name denotations under the (complete) boolean operations is the denotation set for all quantified NPs (*No/Most/All students,*...).*

**Test for Decreasingness (↓):** All Ps are Qs and X is a Q, therefore X is a P.
'No poet' is ↓: if all Londoners drink stout but no poet drinks stout then no poet is a Londoner
Negation is ↓: if Londoner → drinking stout then not drinking stout → not being a Londoner

**Gen 6** *The major ways of building NPs from NPs preserve or reverse monotonicity:*

*a.* *Conjunctions and disjunctions of ↑ NPs are ↑; analogously for ↓ NPs.*
*b.* *Possessive NPs have the monotonicity value of the possessor: X's doctor is ↑(↓) if X is.*
*c.* *Negation reverses monotonicity: not more than two boys is ↓ since more than two boys is ↑*

**query** Which NPs occur naturally in partitives, as in *Two of _*?
yes: Two of *those cats*, two of *John's/the ten/John's ten/my cats*
no: *two of *most cats*, *two of *no cats*, *two of *more male than female cats*

**Gen 7** *Post-*of *DPs of the form Det + Noun denote proper principal filters (= for some $p > 0$, $F(q) =$ True iff $p \leq q$).*

**query** Which NPs occur naturally in Existential There (ET) contexts, as in:
Aren't there *at most four undergraduate students* in your logic class
Weren't there *more students than teachers arrested* at the demonstration?
Just *how many students* were there at the party?
Aren't there *as many male as female students* in the class?
*There are *most students* in my logic class
*Isn't there *the student who objects to that*?
*Isn't there *every student who gave a talk at the conference*?
*Was there *neither student* arrested at the demonstration?

**Gen 8** *Just NPs built from intersective Dets and their boolean compounds (modulo pragmatic factors) occur in ET contexts.*

**Intersective (Generalized Existential) Dets** are ones whose values at a pair $A$,$B$ of properties just depends on $A \cap B$. Formally, they satisfy (3):

(3) $D(A)(B) = D(X)(Y)$ whenever $A \cap B = X \cap Y$.

**some intersective Dets** some, a/an, no, several, more than six, at least / exactly / fewer than / at most six, between six and ten, just finitely many, infinitely many, about / nearly / approximately a hundred, a couple of dozen, practically no, not more than ten, at least two and not more than ten, either fewer than five or else more than twenty, that many, How many?, Which?, more male than female, just as many male as female, no... but John

**Co-intersective Dets** *every*, *all but two*,... which satisfy (3) with $-$ for $\cap$, are not intersective. Nor are **proportionality Dets**: most, less than half, seven out of ten

### 3. Properties of PL not present in NL

**Precision** NL, not PL, is structurally ambiguous

1. John didn't leave because the children were crying
   R1: That's why he stayed [not leave][because the children were crying]
   R2. He left for some other reason [not [leave because the children were crying]]
   **Compare** in PL: $\neg(P \rightarrow Q)$ versus $(\neg P \rightarrow Q)$

2. Every student read a Shakespeare play (over the vacation)
   R1: For every student there was a play he read—maybe different students reaD different plays
   R2: There was one Sh. play that every student read (maybe Hamlet, maybe Lear,...)
   **Compare** in PL: $\forall x \exists y(x < y)$ vs $\exists y \forall x(x < y)$ They have different truth values

3. John told Bill that he had the flu. John said: "I have the flu", "You have the flu", or Henry (identified in context) has the flu. Compare: john$_x$($x$ told bill that $x$ had the flu), bill$_y$(john told $y$ that $y$ had the flu), john told bill that $z$ had the flu.

4. John thinks he's clever and so does Bill [think that John is clever, think that he himself is clever]
   John$_x$($x$ think $x$ is clever & Bill think that $x$ is clever)
   John$_x$($x$ think $x$ is clever) & Bill$_y$($y$ think that $y$ is clever)

**Fact:** NL lacks the variable binding operators of PL.

### 4. Logical resources of NLs not present in PL

**NL quantifiers** take pairs of properties as arguments, the first restricting the domain of quantification, as in *Most **poets** daydream*. PL quantifiers have just one property argument:

a. Some poets daydream = $\exists x(P(x)\&D(x)) \equiv SOME(\lambda x(P(x)\&D(x))$
b. All poets daydream = $\forall x(P(x) \rightarrow D(x)) \equiv ALL(\lambda x(P(x) \rightarrow D(x))$

**Theorem (Keenan 1992)** The domain eliminable NL quantifiers are just the (co)-intersective ones, thus excluding the proportionality Dets.

**Gen 9** *All PL quantifiers are domain reducible; not so in NL.*

**Def** If Det is proportional then the truth of *Det poets daydream* depends on the proportion of poets that daydream. ($DAB = DXY$ whenever $|A \cap B| \,/\, |A| = |X \cap Y| \,/\, |X|$)

Examples: most, seven out of ten, less than half, not one… in ten
*Most poets daydream* does not mean either (For most objects $x$ (Poet($x$) & Daydream($x$)) or (For most objects $x$, if Poet($x$) then Daydream($x$)). BUT

**Gen 10** *a. NL Quantifiers are domain independent: Blik defined by* $BLIK(A)(B) =$ *T iff* $|\neg A| = 2$ *is not a possible English determiner.* Blik cats are black *would be true iff the number of non-cats is two.*
*b. NL Qs are overwhelmingly* ***conservative****: Det As are Bs cannot depend on Bs which are not As, so* $DAB = D(A)(A \cap B)$ *NB: Conservativity (CONS) and Domain Independence (DI) are independent. (BLIK is CONS but not DI;* $F$ *in* $FAB =$ *T iff* $|A| = |B|$ *is DI but not CONS)*

**Gen 11** *Proportionality Quantifiers determine novel reasoning paradigms:*
*Exactly half the students passed. Therefore, Exactly half the students didn't pass.*
*Between a third and two thirds of the students passed the exam. Therefore, between a third and two thirds of the students didn't pass the exam.*
$Qx\phi$ *never entails* $Qx\neg\phi$*, for* $Q =$*'all' or 'some'*

**Gen 12** *Non-trivial Proportionality quantifiers are "logical" (= their denotations are permutation invariant) but* ***not definable*** *in PL. Similarly with cardinal comparatives, of type* $((1,1),1)$*:*
***More*** *poets* ***than*** *priests daydream;* ***Fewer*** *boys* ***than*** *girls,* ***More than twice as many*** *girls* ***as*** *boys;* ***Half again as many*** *girls* ***as*** *boys. These quantifiers may have multiple occurrences:* ***Fewer*** *boys* ***than*** *girls read* ***more*** *poems* ***than*** *plays. Jack read* ***more poems than Jill. A certain number of*** *students applied for* ***a smaller number*** *of scholarships.*

**Gen 13** *PL quantifiers are* ***extensional****, NL ones may not be. In a situation in which the doctors and the lawyers are the same individuals,* Every doctor attended (the meeting) *and* every lawyer attended… *have the same truth value, but* Not enough doctors attended *and* not enough lawyers attended *may have different values. All PL quantifiers are like* every *here.* too many, surprisingly many, … *are like* not enough.

# General Program

# FOCUS AND NEGATIVE CONCORD IN HUNGARIAN

ÁGNES BENDE-FARKAS

IMS
Stuttgart University
agnes@ims.uni-stuttgart.de

**Abstract**

This paper presents a newly disccovered exception to Negative Concord in Hungarian that does not involve double negation. Rather, it is like English *it*-clefts that contain two negative particles that correspond to two instances of negation, in two independent formulae. The paper extends a DRT-based analysis of Hungarian Focus to the simplest cases, mentions some implications for the division of labour between syntax and semantics and suggests a possible method of presupposition accommodation that is required by the more complex cases..

## 1. Introduction

The aims of this paper are (*i*) to present new Hungarian data that show a new kind of exemption from Negative Concord that is problematic to syntax-based, or syntax-driven, theories of Negative Concord (e.g. Zeijlstra 2004);[1] (*ii*) to show how the core cases follow from a DRT-based analysis of Hungarian Focus and (*iii*) to sketch an analysis of the more involved cases. These latter cases are relevant also for ongoing debates on presupposition accommodation in the theoretical literature.

## 2. Negative Concord in Hungarian; The Exceptions

This section reviews the standard facts about negation and Negative Concord in Hungarian, and introduces the exceptions to Negative Concord.

### 2.1. Negation and Negative Concord

Hungarian is a Negative Concord language, as shown in (1): two so-called *n*-words, *senki, semmi* ('no-one' and 'nothing') and the negative particles *nem* or *sem* contribute to *one* negation in the sentence. According to linguists working on negation and Negative Concord (cf. Giannakidou 2002, Zeijlstra 2004) Hungarian is a so-called strong Negative Concord language.[2] The relevant contrast is between Hungarian (1b) and Italian(2b). In simple descriptive terms, in Hungarian the preverbal

[1] Some of the data can be found in an unpubslihed UCLA ms by Anna Szabolcsi; what is new here is the realisation of their significance for Negative Concord.

[2] But see Surányi 2002 for the claim that Hungarian is in fact a hybrid. This claim is based on data different from the data presented here.

position of the *n*-word *senki* 'no-one' is not sufficient to prohibit the occurrence of the negative particle *nem* or *sem* — unlike the Italian example (2b), where the pre-verbal *n*-word *nessuno* 'no-one' blocks the negative particle *non*.

(1) a. Nem látott senki semmi-t
Not saw no-one nothing-Acc
'No-one saw anything'

b. Senki ∗(nem/sem) látott semmi-t
No-one not/SEM saw nothing-Acc
– same –

(2) a. Gianni *(non) ha telefonato a nessuno

b. Nessuno ( *non) ha telefonato

### 2.2. Two Positions for *Nem*

The placement of the negative particle *nem* looks standard at first sight: from (1) one could conclude that it is to be immediately preverbal. The problem is that *nem* can surface in not one but two preverbal positions: in an immediately preverbal position, as in (3a), or in a position that immediately precedes Focus, as in (3b).

(3) a. [János]$_F$ nem látta Marit
'It was John who didn't see Mary'

b. Nem [János]$_F$ látta Marit (hanem Péter)
'It was not John who saw Mary (it was Peter)'

There is consensus in the literature that immediately preverbal *nem* projects *NegP*. The status of pre-Focus *nem* is less clear. (See Olsvay 2000 or Surányi 2002 for discussion.[3])

### 2.3. Exemption from Negative Concord

The two positions for *nem* in Hungarian are relevant for the study of Negative Concord, because, as it turns out, a sentence can have **two** occurrences of *nem*.

(4) Nem [János]$_F$ nem látta Marit (hanem Péter/hanem Péter ÉS János)
'It was not John who did not see Mary (it was Peter/it was Peter AND John)'

It is important to note about (4) that an English *it*-cleft is indeed a close paraphrase for it. That is to say, this is not a double negation sentence.[4]

It is as if (4) had a complex underlying structure, and the two negative particles contributed to different 'compartments' of this structure. This may have motivated, I think, Surányi's proposal concerning the metalinguistic status of pre-Focus *nem*. Sentences like (4) will be said to involve iterated negation, or 'independent' negation.

The sentences in (5) show that *n*-words are licensed in such environments, as expected:

[3] According to Surányi, the pre-Focus particle corresponds to meta-linguistic negation; his own examples are attractive, but some examples later in this paper complicate matters considerably.

[4] (4) does not entail that John saw Mary: it is compatible with a scenario where Mary was not seen by Peter and John. That is, what (4) negates/denies is that John is the unique (maximal) individual with the property of not having seen Mary.

(5) a. Nem [Marit]$_F$ nem látta senki (hanem Katit)
'It was not Mary who wasn't seen by anyone (but Cathy)'
b. Senki nem [Marit]$_F$ nem látta (hanem Katit)
'No member $x$ from a contextually salient group
is such that it was Mary whom $x$ did not see (it was Cathy)'

## 3. Focus, Presupposition, Negation

This section outlines a DRT-based (Genabith et al. ta) analysis of Hungarian sentences that involve Focus and negation. It takes a proper semantic analysis of negation and Negative Concord for granted, as its main concern is the interaction between the presupposition–assertion structure triggered by Hungarian Focus and negation. The scope interactions between presuppositions and *n*-words will be of particular interest, since the Hungarian data provide interesting evidence on presupposition accommodation and domain restriction.

### 3.1. An Analysis of Hungarian Focus; Simple Cases of Iterated Negation

Hungarian Focus is taken to have the following properties (Bende-Farkas 2005) that are of relevance to this paper: (*i*) It is exhaustive, or maximal (viz the denotation of the focussed constituent is maximal relative to a — possibly complex — predicate constructed from the rest of the sentence). (*ii*) It triggers an existence presupposition.[5] (*iii*) Exhaustivity is also taken to have the status of a presupposition. Exhaustivity, or maximality, is relative to a set of alternatives — that is, one property shared by English and Hungarian Focus is that they invoke a set of contextually salient alternatives.

The simplified analysis adopted here builds on Hans Kamp's DRT-based representation of English Focus (Kamp 2004) and my previous work on Hungarian Focus (Bende-Farkas 2002; Bende-Farkas 2005, also in DRT). The simple sentence (6) would, for instance, be represented as (7) (reverting to a linear, flat format, where $\partial$ marks presuppositions).

(6) János [Marit]$_F$ szereti 'It is Mary whom John loves'

(7) introduces a presupposition about context set $C'$, s.t. the Focus variable $\beta$ and the discourse referent of the focussed constituent both have the property of being in $C'$. There is an existence presupposition (constructed from material to the right of Focus): in this case it involves a state $s$ of John loving some $\beta$. $\beta$ is maximal relative to this property (originally maximality has been encoded with DRT's abstraction operator $\Sigma$, but any suitable maximality operator could be used instead). The assertion part simply identifies $\beta$ with Mary.

[5] Except for $MON \downarrow$ *XP*s that are focussed 'by default'.

(7) $\partial(\exists C', \alpha.[C'(m) \wedge C'(\alpha) \wedge m\#\alpha]) \wedge$
$\partial(\exists \beta, s.[C'(\beta) \wedge n \subseteq s \wedge \textit{love}(s)(j, \beta) \wedge MAX(\beta, \lambda\beta'.[\exists s'[\ldots]])]) \wedge$
$\beta = m$

In what follows an abbreviated notation will be used: in (8), the abbreviated version of (7), $\exists\beta^{C'}_{MAX} \ldots$ says that there is a $\beta$ in context set $C'$, and it is maximal (relative to the property that can be 'constructed' from material in the scope of the existential quantifier).

(8) $\partial(\exists\beta^{C'}_{MAX}.\exists s.[n \subseteq s \wedge \textit{love}(s)(j, \beta)]) \wedge \beta = m$

Given the format exemplified in (7), the simplest cases of exceptions from Negative Concord are easy to represent. The point is, pre-Focus *nem* will contribute to the assertion part, and immediately preverbal *nem* will contribute to the presupposition part (it will have narrow scope relative to the 'main' existential quantifier of the existence presupposition). For reasons of space, an explicit representation of such a case will be left as an (easy) exercise. Indeed, iterated negation sentences (without *n*-words) in Hungarian do not present technical difficulties. The problems they present are theoretical. For reasons of space, I will merely enounce these problems here:

A problem concerning syntax is that the 'domain' of NC in these examples is not a syntactic clause, as in most known cases of NC, but a semantic unit: negation in the presupposition triggered by Focus is independent from negation in the assertion part.

Another problem concerns precisely the relationship between presupposition and assertion: in these core cases negation does not have the 'crossing' dependency property that for instance existential quantifiers have (meaning that an existential quantifier in one slot can bind a variable in the other slot).

### 3.2. The Complex Cases: Focus and *N*-words

In general, the scope of the presuppositions triggered by Hungarian Focus conforms to the so-called Scope-Principle (cf. the work of É.Kiss or Szabolcsi), in that linear order corresponds to scope order. (9)–(10) are a minimal scopal pair, for instance: (9) presupposes that there is a unique individual loved by every student, whereas in (10) the presupposition is outscoped by the universal quantifier *minden diák* 'every student'.

(9) [Marit]$_F$ szereti minden diák 'It is Mary whom every student loves'

(10) Minden diák [Marit]$_F$ szereti 'For every student $x$, it is Mary whom $x$ loves

*N*-words also participate in such scope interactions (irrespective of whether the sentence has one or two negative particles). In fact, *n*-words and more standard quantifiers in Hungarian have a uniform behaviour not only as regards the scope of

the presuppositions of Focus, but regarding accommodation problems as well. That is to say, the remarks on accommodation in this subsection will hold not only for Focus and *n*-words but for Focus and other quantifiers as well.

(11) (Nem) $[\text{Marit}]_F$ nem látta senki 'It is (not) Mary who was seen by no-one'

(12) Senki nem $[\text{Marit}]_F$ látta
'For no $x$ (from a given group): it is not Mary whom $x$ saw'

In (11) the presupposition of Focus has wide scope: it is presupposed that there is a unique individual that no-one saw; it is then asserted that this individual is (or is not) Mary. In (12) the presupposition has narrow scope. Assuming that the *n-word* +negation complex translates as *all...not*,[6] the sentence says that for all members $x$ of some group $X$ it is presupposed that there is a unique person that $x$ saw; the assertion is that that person is not Mary.

(13) $\forall x.[\textit{person}(x) \wedge C(x) \rightarrow \partial(\exists\beta^{C'}_{x,MAX}, e, t.[e \subseteq t \prec n$
$\wedge\ \textit{see}(e)(x, \beta_x)]) \wedge \beta_x \neq m]$

Note that the preverbal *n*-word in (12) (and to some extent the post-verbal *n*-word in (11) as well) corresponds to a quantifier with a non-empty domain. In fact, it corresponds to a quantifier with a domain salient in, or familiar from, previous discourse.[7] In my opinion, the relevant factor in this group-specific reading has to do first of all with Focus and the presupposition it triggers. The point is that 'officially' in (12) the presupposition is 'located' in the nuclear scope, but the over-all native speaker impression is that more is presupposed, viz that the presupposition outscopes the quantifer ('for all $x \in X$: there is a $\beta_x$ that $x$ saw'). This is clearest when the sentence has a denial reading, because then there is an obvious antecedent for it. But a sentence like (12) can be uttered without any prior context and a co-operative listener can accommodate it. The question is what gets accommodated, and how.

Simply exporting the presupposition outside the scope of the *n*-word and its associated universal quantifier is not an option, since the presupposition contains a variable bound by that quantifier. Simple intermediate accommodation is not an option, either; it is rejected on the basis of native speaker judgements (that could be further tested).

What does get accommodated, I think, is the proposition that for every member $x$ of some group $X$ there is a $\beta_x$ that $x$ saw. (14) shows the 'antecedent' for the narrow scope presupposition in (13). Taking (13) and (14) as premises, a sequence of

[6] Cf. Giannakidou 2002; Puskás 2000; Surányi 2002

[7] This is indicated by the marked oddness of continuations that state that in fact there was no-one in the intended group. A better example in this respect is the Hungarian counterpart of the following: *Nothing was bought by [Mary]$_F$, because there WAS nothing to buy.* The Hungarian version of this sentence is definitely odd; precisely, I think, because it presupposes a collection of things, and that each of these was bought by someone.

standard inference steps (universal instantiation and inference rules involving conditionals) yield a conclusion that (for all practical purposes) amounts to intermediate accommodation.

(14) $\exists X'.\forall x'.[x' \in X' \rightarrow \exists \beta'_{x'}, e', t'.[e' \subseteq t' \prec n \wedge \textit{see}(e')(x', \beta'_{x'})]]$

**Acknowledgements**

The main observation in this paper was made during a lecture by Arnim von Stechow on features in semantic theory that included Negative Concord as a case study (von Stechow 2005). I thank him, Hans Kamp and Ede Zimmermann for inspiring chats and suggestions — with the usual disclaimers.

**Bibliography**

Bende-Farkas, Á.: 2002, *Comparing English and Hungarian Focus*, presented at Sinn und Bedeutung VII

Bende-Farkas, Á.: 2005, *Maximality: Presupposed and Asserted*, Paper presented at the Szklarska Poręba Workshop on Pragmatics

Genabith, J., Kamp, H., and Reyle, U.: t.a., Discourse Representation Theory, in D. Gabbay and F. Günthner (eds.), *Handbook of Philosophical Logic*, Vol. 13, Kluwer, 2nd edition

Giannakidou, A.: 2002, N-words and Negative Concord, in *The Linguistics Companion*, John Benjamins, Amsterdam

Kamp, H.: 2004, Information Structure in a Dynamic Theory of Meaning, in *Proceedings of the Linguistic Society of Korea*, LSK, Seoul

Olsvay, C.: 2000, A Syntactic Analysis of Negative Universal Quantifiers in Hungarian, *Master's thesis*, TLP Budapest

Puskás, G.: 2000, *Word Order in Hungarian. The Syntax of A-bar Positions*, John Benjamins, Amsterdam

Surányi, B.: 2002, *Multiple Operator Movements in Hungarian*, Ph.D. thesis, Utrecht, LOT

von Stechow, A.: 2005, *Semantic Licensing of Some Uninterpretable Features*, Paper presented at Stuttgart Graduate Seminar, June 9th 2005

Zeijlstra, H.: 2004, *Sentential Negation and Negative Concord*, Ph.D. thesis, Utrecht, LOT

# DYNAMIC SITUATIONS: ACCOUNTING FOR DOWTY'S INERTIA NOTION USING DYNAMIC SEMANTICS

IDO BEN-ZVI

Program of Cognitive Studies of the Language and its Usage

Tel-Aviv University
idobenz@post.tau.ac.il

In this paper Dowty's notion of inertia is further reduced. This is done by seeing normality as operating within a limited context defined by the current conceptually salient vocabulary. Situations are defined as sets of possibilities indiscernible under the vocabulary. Then, using a partial order of eventiveness, normality is given as a test operator on the set of minimally eventive situations in the information state.

## 1. The notion of inertia

The notion of inertia was first incorporated into the semantics of the progressive by Dowty 1979, here reproduced in (1). The guiding intuition behind its use is that the progressive commits the speaker to the eventual completion of the ongoing event if and only if nothing out of the ordinary happens. Enforcing this intuition is the semi-formal definition of the inertia set for a world *w* and interval *I*, *Inr(<I,w>)*. This set is said to include an arbitrary world *w'* iff a.(identity) it is exactly like *w* at all times preceding and including *I* and b. (normality) given the past history of *w*, *w'* is a world in which nothing unexpected happens from *I* onwards.

(1) [PROG $\Phi$] is true at $<I,w>$ iff for some interval *I'* such that $I \subset I'$ and *I* is not a final subinterval for *I'*, and for all *w'* such that $w' \in Inr(<I,w>)$, $\Phi$ is true at $<I',w'>$.

Yet even with the semi-formal constraints at hand, Dowty's analysis of inertia is incomplete. The semantics still presupposes a notion of normality, of a course of events where nothing unexpected happens. Being unable to provide a formal definition to normality, Dowty 'reluctantly concludes' that inertia as a whole must be accepted as a primitive in the semantics. This paper sets out to extend Dowty's theory 'inwards', so to speak, by providing a formal analysis of normality.[1]

## 2. Epistemic considerations of normality

Unexpectedness being a crucial ingredient in the notion of normality, it would

[1] The ideas offered in this paper are articulated at length in Ben-Zvi 2005

seem that no formal account of it is likely to succeed. For how can what is unexpected be brought under the rule of logic? To overcome this hurdle the following mental observation is pointed out. When one speculates about the possible developments of a state of affairs, one does not ponder every logically possible continuation. Rather, a limited set of relevant, or conceptually salient, continuations is envisioned. This is also true of other commonsensical reasoning processes. We use the finite conceptual vocabulary that is available to us at the time and envision only the limited set of continuations that are expressible using that vocabulary. Under this observation normality is seen to concern only events that can be described using the available conceptual context. Within this context, there is a body of sentences that we deem true. They form our knowledge, describing what is *expected*. Complementing this set are the salient descriptions that are merely speculative. These describe the *unexpected*. The possible continuations that we ponder differ by the speculative descriptions that are actualized in them. The normal continuations are those in which the actualization of unexpected event descriptions is reduced to a minimum. As one is treading dangerous ground with these investigations, let us continue immediately to the formalization of these ideas. This is done by extending the dynamic semantics of Groenendijk, Stokhof and Veltman 1996 with a formal apparatus of situations.

## 3. Extending dynamic semantics with situations

The conceptual vocabulary is formalized in (2) as the *linguistic context*. A set of sentences closely related to the discourse through a *growth pattern*. It is assumed that the linguistic context is somehow derived from the actual discourse utterances, and so the context associated with $s[\varphi]$ is formed by applying the growth pattern to the context at $s$ and the sentence $\varphi$. The exact contents of the growth pattern function remain unanalyzed in this paper, except for the two characteristics shown in (3): it is monotone, and it always includes the uttered sentence in its range.

(2) $\Gamma$, the set of linguistic contexts, is the set: $\Gamma = Pow(FORMULAS)$

(3) Function $f \in \Gamma \times FORMULAS \rightarrow \Gamma$ is a *growth pattern* only if for every $\gamma \in \Gamma$ and $\varphi \in FORMULAS$, $f(\gamma, \varphi) \supseteq \gamma \cup \{\varphi\}$

In dynamic semantics the information state can only be divided into the *logical possibilia* of which it consists (the *possibility* structures). To get conceptually salient 'continuations', we need to carve it up more crudely, along the lines set out by the linguistic context sentences. The resulting structures will form *conceptual possibilia*, or *situations*. Each situation is a set of possibility structures (possibilities henceforth) to which the same conceptually salient descriptions apply. Another way of putting it is that each situation is comprised of conceptually indiscernible possibilities. Formally, we start out in (4) by defining the consistency relation between possibilities and sentences, based on the dynamic semantics relations of *possibility extension* and *possibility similarity*. Now (5) defines

indiscernibility between possibilities, relative to a linguistic context. Two possibilities are indiscernible when they are consistent with the exact same set of linguistic context sentences.

(4) Let $i \in I$ (the set of possibilities in dynamic semantics); $\varphi \in FORMULAS$.
$\varphi$ is *consistent* with *i, Cons(i,φ),* iff
$\forall s \in S \; s.t \; i \in s : \exists i^* \in s[\varphi] : i^*$ is similar to an extension $i'$ of *i*.

(5) Let $i,i' \in I$; $\gamma \in \Gamma$. *i* and *i'* are *indiscernible* in $\gamma$ iff
$\forall \varphi \in \gamma : Cons(i, \varphi) \leftrightarrow Cons(i', \varphi)$ .

The definition of situations, in (6), is a little cumbersome due to them containing not only a non empty set of possibilities, but also the linguistic context under which these possibilities are indiscernible. Also, the free variables in the context must be defined in the referent set shared by the possibilities. Lastly, the set must be maximal in the sense that no indiscernible possibility be left out of it.

(6) *M*, the set of situations, is the set

$$\left\{ \left. \begin{array}{l} \gamma \in \Gamma, J \subseteq I, J \neq \varnothing, \\ \forall i, i' \in J : \; i \text{ and } i' \text{ share referent system and are indiscernible in } \gamma, \\ Dom(r) \supseteq FREE\text{-}VARS(\gamma), \\ \text{For every } i \in J \text{ and } i' \in I, \text{ if } i \text{ and } i' \text{ share referent system and world} \\ \text{and are indiscernible in } \gamma \text{ then } i' \in J \end{array} \right| < J, \gamma > \right\}$$

The partiality of situations is made manifest in (7), with their denotation function. A sentence is true (false) in a situation only inasmuch as the sentence (or its negation) is consistent in all of the underlying possibilities.

(7) Let $\varphi \in FORMULAS$, $m \in M$, $m=<J,\gamma>$. The denotation of $\varphi$ in situation *m*, $[\![\varphi]\!]_m$ ,is defined as follows: $[\![\varphi]\!]_m = \begin{cases} 1 & if \quad \forall i \in J : cons(i, \varphi) \\ 0 & if \quad \forall i \in J : cons(i, \neg \varphi) \\ undefined & otherwise \end{cases}$

To wrap up the introduction of situations into the framework, information states are redefined in (8) as sets of situations which share the same linguistic context, under which they are discernible from each other.

(8) *S* ,the set of *information states* (based on situations), is the set

$$\left\{ \left. \begin{array}{l} \forall m, m' \in s : \\ m \text{ and } m' \text{ have the same referent system and} \\ \text{linguistic context, and are discernible this context} \end{array} \right| s \subseteq M \right\}$$

It turns out that the situations in an information state form a partition on the underlying set of possibilities, based on the indiscernibility relation defined by the linguistic context. In accordance with the new structures, the object language semantics are also renovated by using situations in place of possibilities (but without making any other change), and the update function on information states is

redefined as a two stage process. To update state *s* with sentence $\varphi$, first the linguistic context is updated by applying the growth pattern to the current context and the sentence $\varphi$. This causes a repartitioning of the existing situations, but does not affect the underlying population of possibilities.[2] Only then does the standard semantic interpretation of $\varphi$ take place (based on situations in place of possibilities). Even though the object language is now redefined based on situations, the updated framework can be shown to be isomorphic to the original framework under the update function, for as long as no new object language operators are introduced. This is stated more formally in (9).

(9) FACT: Let *S* and *T* be the sets of information states in the original and updated dynamic semantics respectively. With $[]_S$ and $[]_T$ their respective update functions.

There exists a function $F : S \xleftrightarrow{1\text{–}1} T$ in which the following holds:
If $s \in S$, $s=s_0[\varphi_1]\,[\varphi_2]...\,[\varphi_n]$ where $s_0$ is an initial state, and where $\{\varphi_i\}_{i=1..n} \subset FORMULAS$, then $\forall \varphi \in FORMULAS : F(s[\varphi]_S) = F(s)[\varphi]_T$

## 4. Normality again

We now return to normality. Seeing that in this paper events replace intervals, a trace function $\tau$ from events to their temporal intervals is used to maintain temporal ordering. Moreover we assume, for simplicity's sake only, a single domain of entities that contains both objects and events. These entities can be quantified over in the object language.

By defining a partial order of *eventiveness* on the situations in the information state and then selecting the minimal elements, we get the 'least unexpected', or normal, situations. One situation is less eventive than another one if for every conceptually expressible eventive fact that holds in the former situation, a similarly described fact also holds in the latter one. Conceptually expressible eventive facts are facts which are described by linguistic context sentences asserting the existence of an event, i.e. sentences such as $\exists x Event(x) \wedge Walk(x) \wedge Ag(Mary, x)$. We may safely ignore the expected/unexpected distinction while ordering, as expected descriptions are uniformly actualized in all situations.

This complex comparison is set out formally in (10). The required event descriptions are sought out by iterating first on the linguistic context sentences and then on the active quantifiers they define.[3] Filtering out quantifiers that don't represent events is done by appending '$\wedge event(q)$' to every sentence.

---

[2] Note that as the linguistic context monotonically grows, repartitioning after updating the context can only result in a finer grained partition.

[3] The syntactic function *Aq(φ)*, introduced in Groenendijk and Stokhof 1991, gives the active quantifiers of $\varphi$. These are the existential quantifiers introduced in $\varphi$ whose scope is not bound within the sentence.

*Let* $s \in S; m, m' \in s$: *m is less eventive than m' in s, iff*

(10) $\forall \varphi \in \gamma$ *the linguistic context of situations in* $s, \forall q \in Aq(\varphi)$:

$[\![\varphi \wedge event(q)]\!]_m = 1 \Rightarrow [\![\varphi \wedge event(q)]\!]_{m'} = 1$

A new object language operator, *normally*, is now defined as a test in (11). The operator makes use of the minimal elements in the ordering (labeled as *least eventive*) to check if a given sentence (the progressive event's full description as we will see) is indeed valid in every normal situation

$$(11)\, s[normally(\varphi)] = \begin{cases} s & if\ \ \forall m \in s : LeastEventive(m) \rightarrow [\![\varphi]\!]_m = 1 \\ \varnothing & otherwise \end{cases}$$

## 5. The rest of the semantics

Proceeding in brief through the rest of Dowty's semantics, we first take note that the identity requirement in inertia is automatically fulfilled by dynamic semantics. The information state only contains possibilities, and therefore also situations, that are consistent with the knowledge gained thus far. In addition, it is assumed that the reference time interval (denoted by Dowty as *I* but here as *RT*) is at our disposal when we come to analyze the progressive. Finally, for activities the subinterval property must also be postulated.

Given that sentence $\Phi$ asserts the existence of an event described by $\varphi$, that is – $\Phi$ is of the form $\exists e\big(event(e) \wedge \varphi(e)\big)$, the semantics of *PROG(Φ)* is given in (12) as two consecutive updates. The first update asserts what must already be known: that there is an event going on relative to reference time. The second update consists only of the test operator *normally*, that checks if every normal world actualizes the verb description as a fact.

$$(12)\ s[PROG(\phi)] = s\left[\exists e\begin{pmatrix} event(e) \wedge RT \subset \tau(e) \wedge \\ RT \text{ not a final subinterval of } \tau(e) \end{pmatrix}\right]\left[normally(\varphi)\right]$$

## 6. Example

Suppose that having started in a state of ignorance, we now see John heading toward the other side of the road. Accordingly we update our information state with (13), the informal notation being used to keep things as simple as possible. Suppose further that the growth pattern is such that along with (13) it extends the linguistic context with the speculative descriptions in (14) during the update.

(13) $\exists e_{JohnWalkingDirOtherSideRoad} \wedge \tau(e) \supset RT$

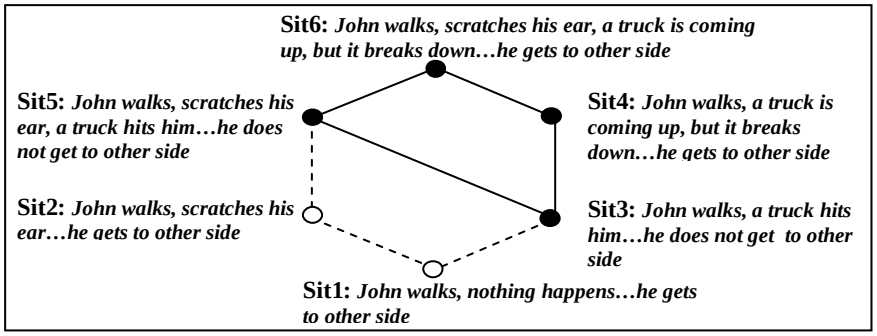
(14) a. $JohnCrossRoad(e)$ c. $\exists e''_{TruckFlyingDownTheRoad} \wedge \tau(e'') > RT$

b. $\exists e'_{JohnScratchHisEar} \wedge \tau(e') > RT$ d. $\exists e'''_{TruckBreakDown} \wedge \tau(e''') > RT$

This information state may even be shared between us as bystanders and John as the agent. Except that, our darker fears materializing, the speeding truck now

makes an appearance. Updating our information state (but not John's, he is oblivious to the truck), (14c) now gains the status of an utterance. Graph (15) maps the situations in the two information states according to their eventiveness. The ordering is shown by the connecting arches, with less eventive situations always lower than more eventive ones. John's information state is displayed by the whole graph, while ours is just the solid parts of it.

(15) **John's attempted crossing of the road**




(16) John is crossing the road

(17) John was crossing the road when the truck hit him.

Evaluating the truth of the progressive sentence (16), we see the two information states each provide a different situation as normal. For John it is situation 1, and the sentence true. For us though situation 3 is minimal, and the sentence is false.

To conclude, it is interesting to point out that sentences such as (17) are not accounted for in the suggested framework, for they require simultaneous use of both information states to come out true. Extending dynamic semantics with multiple concurrent information states may provide an answer to such sentences.

**Acknowledgements**

I am greatly indebted to Fred Landman for helping throughout, and to Nissim Francez for getting me to write this paper.

**Bibliography**

Ben-Zvi, I.: 2005, *Dynamic Situations: Accounting for Dowty's Inertia Notion Using Dynamic Semantics (Thesis 2005)*, http://www.semanticsarchive.net

Dowty, D.: 1979, *Word Meaning and Montague Grammar*, Kluwer, Dordrecht.

Groenendijk, J. and Stokhof, M.: 1991, Dynamic Predicate Logic, *Linguistics and Philosophy* 14: 39-100.

Groenendijk, J., Stokhof M. & Veltman F.: 1996, Coreference and Modality, in: S. Lappin (ed), *The Handbook of Contemporary Semantic Theory*, pp 179-216, Blackwell, Oxford

# EXHAUSTIVITY, HOMOGENEITY AND DEFINITENESS

RICHARD BREHENY

Dept. of Linguistics UCL
richardb@ling.ucl.ac.uk

In this paper, it will be argued that the Homogeneity Presupposition (Fodor 1970, Loebner 2000, Schwarzschild 1993) does not provide an adequate account of the tendency of plurals to obtain exhaustive, 'any'-interpretations in negative contexts. It is argued that Krifka's (1996) rule for plural predication would do better if it were restricted to arguments which are in some sense definite. An analysis is sketched which locates the optionality of plural interpretations in definite noun phrases rather than in the predication.

## 1. The HP account of 'any' readings

The tendency for definite plurals in (1a,b) to be interpreted exhaustively cannot be accounted for simply in terms of their denoting the maximal set of children:

1. a. Mary saw the children
   b. Mary didn't see the children

In particular, the reading of (1b) according to which Mary doesn't see any of the children is problematic since it cannot always be attributed to scope. In (2) the 'not any' construal is still prominent in spite of the fact that the plural noun phrase could not move/scope above the negative element:

2. $No_i$ woman likes the children in $her_i$ care

A Homogeneity Presupposition (HP) has been proposed to account for these intuitions: A definite plural presupposes that each individual in its denotation behaves the same with regard to the predicate with which it combines. (3) represents an implementation for one-place plural predications:

3. $*P(X) = 1 \text{ iff } \forall x[x \in X \rightarrow P(x)]$
   $0 \text{ iff } \forall x[x \in X \rightarrow \neg P(x)]$ Undefined otherwise

For Schwarzschild (1993) the HP results from the interaction of the closure operator on distributive predicates and his four-valued semantics. For others, such as Loebner (2000), the HP is a more general phenomenon. The latter view seems supported by the fact that the phenomenon is not limited to plural predications. In (4) and (5) below, the (b)-examples are most readily understood to suggest that no part of the relevant activity was undertaken.

4. a. John walked to the store
   b. John didn't walk to the store
5. a. John read the book
   b. John didn't read the book

It is well known that the 'completeness' implications of the (a)-examples are very defeasible - as where 'the book' in (5) is the bible, or where the sentence is modified by 'on his lunch break'. Similarly the HP-effect in the (b) cases.

Given that this is a general and independent phenomenon, it seems a good place to start to explain the problematic 'any' construals of plurals where scope-taking is not possible. However, there are some criticisms that one could make of this account. First, that the effect should be thought of as due to a presupposition of *homogeneity* is questionable. It seems that, at least with 'any' readings of plurals, the effect can be cancelled easily by invoking contextual factors which do not particularly conflict with the idea of homogeneity. In (6a,b), the 'not all' construals are made prominent but context does not conflict with the putative HP:

6. a. No farmer vaccinated the donkeys in his care by the deadline
   b. Mary was charged with contempt of court because she didn't answer the judge's questions.

Moreover, the 'any'-effect can occur where the HP would be implausible. For (7) the 'any' reading is prominent and yet it is not so plausible that the gang members would be caught all-or-none:

7. The gang members dispersed to the four corners of the world after the bank robbery, and the police haven't found them yet.

More seriously, it seems we obtain the 'any' readings in the absence of negation - i.e. where the HP account predicts the 'all' construal. In fact, they become prominent in DE contexts generally - cf (8) which has a prominent 'any'-reading

8. Every/No farmer who abused the donkeys in his care was prosecuted

This problem also emerges for embedded questions. Consider: the 'any' reading is equally available for both (9a) and (9b). If we were to use the HP to derive the reading for (9a), we ought to similarly accommodate the HP for (9b) - but this should result in the reading where Mary wonders whether John ate all the cookies:

9. a. Mary doubts that John ate the cookies
   b. Mary wonders whether John ate the cookies

## 2. Krifka's Alternative

Taking together the problematic data in (6-9) and the idea that the 'any'-effect has something to do with how we understand predications generally, (4-5), it is tempting

to take the following line: Natural language semantics leaves open to some extent what kinds of states of affairs would verify certain predications. By a pragmatic process (complementary to Q-implicature) we tend to draw inferences about the class of situations a given predication is about - effectively coercing the interpretation of the predicate. The effect of the guiding interpretive principle is usually described in terms of the Strongest Meaning Hypothesis (see Dalrymple et al 1998).

Krifka (1996) advocates applying this line of thinking to plural predications:

> "If a predicate P applies to a sum individual x, grammar does not fix whether the predication is universal ($\forall$y[y $\subseteq$ x $\rightarrow$ P(y)]) or rather existential ($\exists$y[y $\subseteq$ x $\wedge$ P(y)]), except if there is explicit information that enforces one or the other interpretation."

This goes together with a version of the SMH. As such we can now predict why the 'any' readings are favoured in DE contexts, and so forth. We also understand the relation between the presence of the favoured readings and contextual factors.

However there are problems with this proposal. Consider that (10) can be understood so that 'two cookies' is both non-specific and narrow scope:

10. No one ate two cookies.

On this kind of reading, the sentence can only be understood in the weaker 'not two' way. There seems to be no conceivable context where it gets understood as 'No one ate any cookies'. And yet this ought to be the favoured understanding according to Krifka's proposal - assuming 'two cookies' quantifies over collections of cardinality two. A similar problem seems to arise with conjunctions. It is difficult to understand (11) to mean that no boy invited either parent:

11. No boy invited his mother and (his) father to the school concert

If we look back to the cases where Krifka's account seems to work, they involve plural definite descriptions (or other definites). The account seems to have come unstuck when we looked at non-definite plural forms in (10) and conjunctions (11).

It is an interesting fact, however, that when 'two N' is specific, the 'any' reading readily re-appears in DE contexts. E.g. where 'two internet sites...' in (12) is specific, quantification is duly construed as being over people who gave to *either* site:

12. No one/Everyone who gave money to two internet sites selling plots on Mars complained to the police

Examples can be multiplied - consider (13) where again binding limits scope-taking:

13. No politician likes to answer two questions about his private life {whether he has inhaled or had an extra-marital affair}

As an aside, it is worth noting that (12) provides a counter-example to the claim by Reinhart 1997 and others that exceptional wide scope specific indefinites cannot receive a distributive, 'any' reading. Geurts (2002) gives others but uses this data to

support a movement account of specifics. A third alternative is offered below.

Regarding conjoined noun phrases, Szabolcsi & Haddicam (2004) note that Krifka's predicted readings are preferred in a number of languages including Hungarian. S&H claim that the 'neither...nor' reading is available for English informants only under certain conditions: either when the individuals are a 'package' (physics and algebra vs hockey and algebra) or where there is some contextual expectation that the individuals 'come together'. Although this characterisation is difficult to square with (11) above (since mothers and fathers ought to be easily thought of as a package) there seem to be genuine 'any' readings with conjunctions. Consider for instance that Philby, Burgess & MacLean were a famous cohort of Cold War double-agents, then (14) can be understood to involve quantification over people who were briefed by any of the three:

14. Everyone who was briefed by Philby, Burgess and MacLean was misinformed

One way to recast the generalisation in S&H is to say that the 'neither...nor' readings of conjunctions are a possibility in English (if not all languages) where the conjunction is understood in a 'definite collection' manner.

## 3. An Alternative Perspective

Given the above considerations, there is reason to doubt that Krifka's line of thinking will work for plural predications. I.e., it seems that (in the non-specific case) for 'John read two books' to be true, John has to have at least dipped into each of a collection of two books. Further, since the problematic 'any' readings seem to emerge only when the noun phrase is definite or specific, it may be that we could look to these noun phrases as the location of the underspecification. The account sketched in the rest of this paper does just that. To this end, it is assumed that descriptions are just existential QNPs requiring contextual restriction like all QNPs. The potential variability is due to the manner in which the presuppositions of plural definites/specifics affect domain restriction of their interpretation.

### 3.1. The interaction of the semantics and pragmatics of descriptions

While it has appeared puzzling why the 'any'-reading emerges in cases like (2), no one is much surprised when the same readings occur for non-definites:

15. No one saw children in the park

But there is a sense in which this is just as problematic for accounts which treat descriptions as (complex) terms denoting individuals - as do choice function accounts. In that case, an otherwise unmotivated local ∃-closure mechanism needs to be invoked to bind the choice function variable in the scope of the negative quantifier. By contrast, the reading of (15) is entirely unsurprising if we assume

indefinite descriptions are existential quantificational phrases. The main motivation for the choice function account over the quantificational account of indefinites comes from extraordinary scope readings of specific indefinites (Reinhart 1997). But there is a good quantificational alternative according to which specifics involve domain restriction down to a condition which the speaker could provide (see Schwarzschild 2002, Breheny 2003). In the singular case, this effectively makes quantification over a domain of one. In the plural case, however, it is an open question whether restriction would always result in a domain of one (the maximal set satisfying the speaker-dependent condition). As already suggested, specific indefinites enter into the 'all'/'any' alternation in, respectively, UE & DE contexts. So, we understand (16a) so that every student looked for all of certain linguists at the conference; while for (16b) no student looks for any of certain linguists:

16. a. Every student$_i$ looked for certain$_i$ linguists at the big conference
    b. No student$_i$ looked for certain$_i$ linguists at the big conference

On a quantificational account of (16a,b), the specific indefinite's domain gets restricted by *certain_ linguist$_u$(i)* which (relative to student *i*) expresses the property of having whatever property of collections the speaker has in mind in uttering the noun phrase, u. Now, suppose that in uttering (16a) I have in mind the idea that each individual student hero-worships one or more linguists. In that case, what I could just as well have in mind is either the condition *λX[*linguist(X) ∧ *worships(X)(i)]* or *max(λX[*linguist(X) ∧ *worships(X)(i)])* (where *max* ≡ *λXλY[Y = U{Z:X(Z)}]*). So there is an underdeterminacy here as to what condition I have in mind - regardless of the idea behind that condition. As such, it could be presupposed either that I have the maximal or non-maximal condition in mind. The choice of implication would in turn be driven by the pragmatic condition which maximises informativity in other SMH cases.

Turning now to definite descriptions, we can ask what makes them different from indefinites. One common way to answer this question is to say that definiteness serves to signal that there is a salient condition (accessible to the audience) by which one can identify the individual or set of individuals 'under discussion'. Let's consider where the condition is made salient by discourse:

17. Some students and some faculty attended the workshop. The students were charged a reduced rate.

Again we can observe that more than one potential restrictor for the definite is made salient by the previous discourse: *λX[*student(X) ∧ *attended_WS(X)]* or *max(λX[*student(X) ∧ *attended_WS(X)])*. Given the attested variability in quantificational force of definites, it seems that these offer equally available restrictions on quantification. One way to think of how this comes about is that the identifying condition on the individuals in question is just *λx[student(x) ∧*

*attended_WS(x)]* but that in restricting the (((et),t),(((et),t),t)) existential determiner of the definite description the maximal singleton restrictor is just one option. The assumption is that the SMH chooses.

It seems reasonable to pursue the idea that in both the definite and the specific case, it is assumed that identifying conditions correspond to expressions of properties of individuals. If so, the uniqueness implication on singulars would come via the number marking which signals that a singleton is under discussion. I.e. where an *identifying* condition is for a singleton, it follows that that condition is only satisfied by one thing. Along similar lines, one can argue that in 'the two students', 'two' contributes to the presuppositions of the description. Suppose that the second segment in (17) above is, 'The two students were charged a reduced rate.'. Then there would be an indicated identifying condition *λx[student(x) ⩘ attended_WS(x)]* and in addition a presupposition that this condition is satisfied by two individuals. Similarly for '(a certain) two students', the number marking arguably contributes to the presuppositions about the collection the speaker has 'in mind'. In both cases, there is underspecification as to how the restrictor of the plural description is fixed leading to the possibility of 'any' readings as in (12-13) above. As for definite conjunctions ('Philby, Burgess and MacLean'), these could be covert plural definite descriptions with the named individuals providing the identifying condition (λx [x = Philby ⩖ x = Burgess ⩖ x = MacLean]).

## Bibliography

Breheny, R.: 2003, Folk ideas about reference and specific indefinites (revised s&b7 paper) (http://www.phon.ucl.ac.uk/home/richardb/talknotes.pdf)

Dalrymple, M., Kanazawa, M. Kim, Y., Mchomobo, S. & Peters, S.: 1998, Reciprocal Expressions and the Concept of Reciprocity, in *L&P* 21: 159–210.

Fodor, J.D.: 1970, *The Linguistic Description of Opaque Contexts*, Ph.D, MIT.

Geurts, B.: 2002, Specific indefinites, presupposition and scope, in R. Bauerle, U. Reyle & T.E. Zimmermann (eds) *Presupposition and Discourse*. OUP, Oxford

Krifka, M.: 1996, Pragmatic strengthening in plural predications and donkey sentences, in T. Galloway & J. Spence (eds.), *SALT VI*. Cornell U. Press.

Löbner, S.: 2000, Polarity in Natural Language: Predication, Quantification & Negation in Particular & Characterizing Sentences, in *L&P* 23(3), 213–308

Reinhart, T.: 1997 Quantifier-Scope: How labour is divided between QR and choice functions, in *Linguistics and Philosophy*, 20:335-397

Schwarzschild, R.: 1994, Plurals, presuppositions and the sources of distributivity, in *Natural Language Semantics* 2, 201–248.

Schwarzschild, R.: 2002, Singleton Indefinites, in *J. of Semantics* 19.3:289-314.

Szabolcsi, A. & Haddicam, B.: 2004, Conjunction meets negation: A study of cross-linguistic variation. *Journal of Semantics* 21(3) 219-250

# COMPLEX ANAPHORS – ONTOLOGY AND RESOLUTION[1]

MANFRED CONSTEN & MAREILE KNEES

Institut für Germanistische Sprachwissenschaft
Friedrich-Schiller-Universität Jena
manfred.consten / mareile.knees @uni-jena.de

We will describe anaphoric complexation processes and their constraints in terms of ontological categories. Furthermore, we will provide a resolution model for complex anaphors based on semantic as well as conceptual structures, thus integrating DRT and cognitive approaches. An example of an ambiguous complex anaphor will be discussed in order to show the role of ontological constraints in complex anaphora processing.

## 1. Introduction

Complex anaphors are nominal expressions referring to propositionally structured referents[2] (such as propositions, states, facts and events) while introducing them as unified entities into a discourse representation. Additionally, they can classify or evaluate the referent.

(1) *Young drivers usually drive too fast. This/ this fact/ this image/ this impertinence …*

Researchers have referred to complex anaphors heterogenously, e.g. *abstract object anaphora* (Asher 1993, 2000) or *situational anaphora* (cf. Fraurud 1992). From a semantic point of view, complex anaphors present DRT-approaches with a challenge, as resolving them involves conceptual knowledge.

## 2. On Complexation Processes

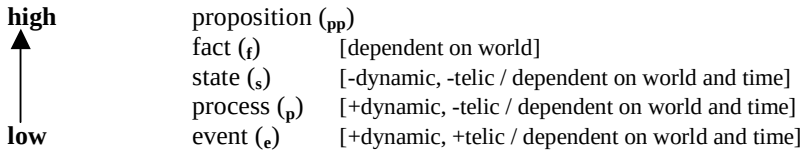
Complex referents are propositionally structured objects, that have been topic of several detailed analyses: There is no final agreement on the ontological categorisation of such

[1] This paper has been written within the context of the research project "KomplexTex", granted by the Deutsche Forschungsgemeinschaft (SCHW 509/6-2). In the framework of this project we have used the TigerCopus in order to systematically determine different grammatical and ontological types of complex anaphors (cf. Consten/ Knees/ Schwarz-Friesel forthc.).

[2] "Let us here use 'referent' for the discourse entity referred to, regardless of its level of representation" (Fraurud 1992: 26). For levels of representation see section 3.

referents as events, states, processes or situations (cf. Asher 1993, 2000, Higginbotham 2000, and Maienborn 2003). Nonetheless, we get the following classification showing the increasing abstractness of the proposed ontological types.

(fig. 1): degree of abstractness — ontological category




### 2.1 Types of Complexation Processes

Now let us have a closer look at the complexation process. We distinguish between three types of complex anaphoric reference (s. (a) – (c)).

(a) The ontological status of the referents stays the same during the anaphoric process, since the antecedent and the anaphor denote the same ontological type ($z_x \approx x$ )[3], s. (2).

(2) [*The Americans tried to invade the building but were forced back by shots from the top floor.*]$_{e}$ *It is said that two soldiers were injured during* [*this action*]$_{e}$, *one inside the house and the other one outside the house.* (TigerKorpus)

(b) The anaphorical expression itself is neutral with respect to ontological types. For this reason, the discourse entity established by the anaphoric process usually keeps the ontological type denoted by the antecedent ($z_{neutral} \approx x$).

(3) [*The Americans tried to invade the building but were forced back by shots from the top floor.*]$_{e}$ [*This*]$_{n}$ *happened yesterday while Mr. Rumsfeld visited Bagdad.*

Even though the anaphor is neutral with respect to ontological types, there are cases where a different type is fixed by the syntactic/semantic context provided by the sentence the anaphor is part of. In (4), the event-referent must be factual in order to serve as a proof.

(4) [*The Americans tried to invade the building but were forced back by shots from the top floor.*]$_{e}$ [*This*]$_{n}$ *proves that the situation isn't under control yet.*

(c) Due to its lexical meaning, the anaphoric expression denotes another ontological type than its antecedent. Thus, the anaphorical process changes the ontological type of the referent ($z_x \approx y$).

[3] '≈' assigns a complex referent (x) to an anaphor (z) (cf. Asher 1993: 145).

(5) [*The Americans tried to invade the building but were forced back by shots from the top floor.*]$_{\mathbf{e}}$ [*This fact*]$_{\mathbf{f}}$ *proves that the situation isn't under control yet.*

In (5) the event referent denoted by the antecedent becomes a fact.

(6) [*Instead of working on her training report, she went out to eat ice cream.*]$_{\mathbf{e}}$ *I won't tolerate* [*this hanging out*]$_{\mathbf{s}}$ *any longer*. (oral communication)

Here the single event (the referent's going out to eat ice cream) is released from its concrete temporal and spatial fixation by the state-anaphor *hanging out*, thus it is understood as a typical, exemplary incident.

## 2.2 Constraints on Ontology Changing Complexation

(7) [*The earth turns about the sun.*]$_{\mathbf{p}}$ [*This process*]$_{\mathbf{p}}$ / [*This state*]$_{\mathbf{s}}$ *will presumably last for* $7 \cdot 10^9$ *years.* [*This fact*]$_{\mathbf{f}}$ *is well known since the Middle Ages. Researchers of the Vatican were not allowed to examine* [*this possibility*]$_{\mathbf{pp}}$ / *[*This event*]$_{\mathbf{e}}$…

As the example shows, anaphorical complexation can shift referents of any ontological type to a discourse entity of either the same ontological type or an ontological type that is more abstract. Thus, anaphorical complexation can be a process of increasing abstractness (s. fig. 1).

(8) $*z_y \approx x$ if x > y ("if x is higher on abstractness scale than y")

This 'abstractness-constraint' can serve to explain ontological based resolution of ambiguous complex anaphors:

(9) [*The Jacobs-Sisters are always in a wonderful mood and flashy.*]$_{\mathbf{s}}$ [*Yesterday they had a great performance in New York.*]$_{\mathbf{e}}$

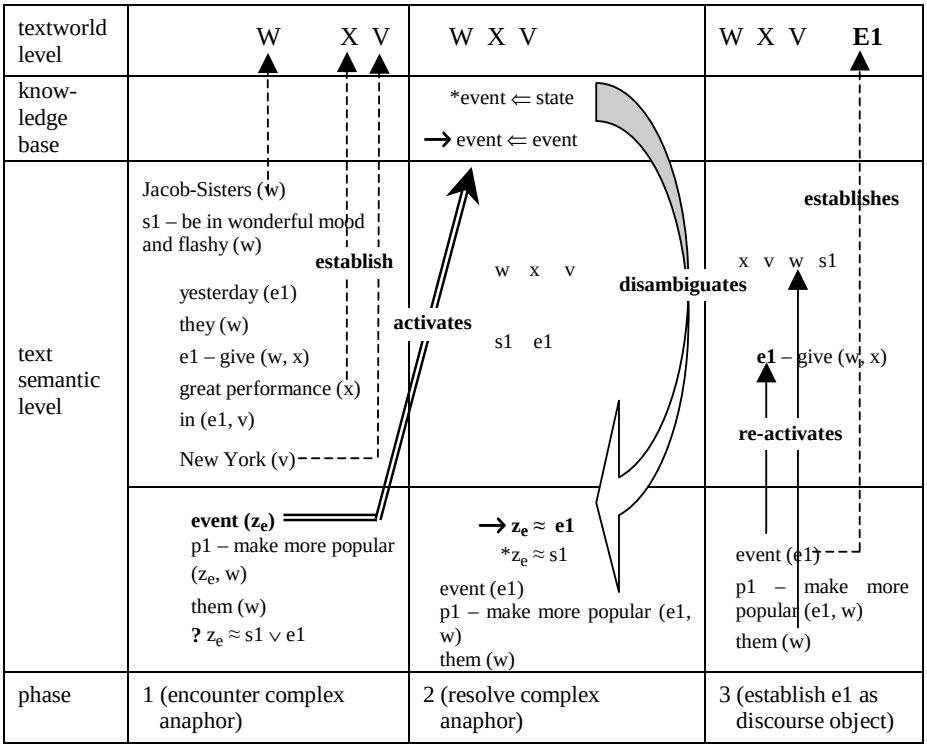
(a) [*This event*]$_{\mathbf{e}}$ *has surely made them even more popular.*
(b) [*This quality*]$_{\mathbf{s}}$ *has surely made them even more popular.*
(c) [*This / that*]$_{\mathbf{n}}$ *has surely made them even more popular.*

The two complex anaphors ((a) vs. (b)) have different antecedents, although both sentences in (9) are accessible as possible antecedents for both of the anaphors from a pure structural point of view (as version (c) shows). However, the first sentence is ruled out as antecedent in case of (a) since an event-anaphor cannot be assigned to a state-antecedent. In case of (b), there is no such restriction (as (6) shows it is possible to assign state-anaphors to event-antecedents in principle) but there seems to be a preference for an antecedent of the same ontological type if provided by the preceding text. These kinds of disambiguation are difficult to explain in terms of purely structural constraints (like DRT-approaches).

## 3. Processing Complex Anaphors

So in our model (taken from Consten / Knees forthc.) we will integrate procedural aspects in using a combination of DRS and cognitive Textworld Models (Schwarz 2001). We distinguish between different levels: the text semantic level, the textworld level and the knowledge base.

(fig. 2) Resolution model for (9a)




Referents are introduced by textual structures at the text semantic level. The nominal expressions *Jacob-Sisters*, *great performance* etc. in example (9a) introduce referents at the text semantic level (w, x... as illustrated in fig. 2).[4] Moreover, as nominal

[4] *They* in the second sentence is immediately resolved to the Jacobs-Sisters since it refers to the only plural

expressions they directly establish discourse entities at the textworld level (W, X…) by activating the corresponding concept in the long term memory (phase 1). The textworld level represents the discourse entities which are talked about in the discourse. In contrast to the nominal expressions, propositional expressions introduce complex referents (like events, states etc.) only into the text semantic level (e1, s1…) but they do not establish discourse entities at the textworld level. The knowledge base contains different sources of knowledge e.g. lexical or conceptual knowledge.[5]

Initially, anaphors do not establish discourse entities at the textworld level but are interpreted at the text semantic level where the appropriate part of the textual structure is re-activated. In case of complex anaphors, these textual parts are propositionally structured. So in example (9a) the complex anaphor z (*this event*) of type e ("event") denotes due to its lexical meaning an event-referent. In phase 2, the anaphor ($z_e$) activates knowledge about ontological categories (i.e. the abstractness-constraint (8) ($z_e \approx$ e1; $*z_e \approx$ s1)) and so the anaphor is assigned to the adequate prementioned referent. In phase 3 the anaphor re-activates this propositionally structured referent and thereby establishes it as a unified discourse entity E1 at the textworld level. Thus, complex anaphors differ from (direct) nominal anaphors as the latter refer to objects already introduced as discourse entities. DRT approaches do not reflect this difference as they assume that each incidence of an anaphor integrates a new discourse referent at the DRS (cf. the critical remarks in Löbner 1985: 320, Cornish 1999: 186, and Consten 2004: 61).

Once the complex referent is established as a unified discourse entity by a complex anaphor, the discourse entity is accessible by personal pronouns (as *it* in the 3rd sentence),[6] whereas the use of personal pronouns in the Vorfeld as a complex anaphor (as *it* in the 2nd sentence) is restricted (cf. Hegarty 2003):[7]

(10) [*The earth turns about the sun.*]$_{\mathbf{p}}$ [*This process*]$_{\mathbf{p}}$ / [*This*]$_{\mathbf{n}}$ / *[*It*] *will presumably last for* $7 \cdot 10^9$ *years.* [*It*] *might, however, terminate a few years earlier .*

---

entities previously introduced into the discourse namely the Jacobs-Sisters.

5 We restrict our illustration to those parts of knowledge that are used in order to resolve the complex anaphor. The preferred interpretation is marked by an arrow in the figure.

6 *It* in the 3rd sentence is not a complex anaphor since it is not assigned to a propositional structured antecedent but to a NP-antecedent (*This hanging out* / *This)* by which a unified discourse entity has already been established.

7 Hegarty (2003: 1-2) assumes that events introduced by a clause are immediately accessible by personal pronouns since they are in focus merely due to their ontological status. However, some of our data does not support his claim. We have no evidence that ontological states of referents are determinants of a salience hierarchy.

## 4. Summary

We have defined complex anaphors as anaphors that condense prementioned propositional referents establishing them as unified discourse entities. Anaphoric complexation is a process of increasing abstractness with respect to ontological categories. Thus, we distinguish between neutral and ontology changing complexation and propose an "abstractness constraint" which serves to explain the resolution of certain kinds of ambiguous complex anaphora not solved by current approaches. Our sketch of a process model of anaphoric complexation is able to integrate cognitive aspects of language processing into a formal semantic framework.

## Bibliography

Asher, N.: 1993, *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer

Asher, N.: 2000, Events, Facts, Propositions and Evolutive Anaphora. In: Higginbotham, J./ Pianesi, F./ Varzi, A. (eds.), *Speaking of Events*. Oxford: University Press, pp.123-150

Consten, M.: 2004, *Anaphorisch oder deiktisch? Zu einem integrativen Modell domänengebundener Referenz*. Tübingen: Niemeyer (LA 484)

Consten, M./ Knees, M.: forthc., Complex Anaphors in Discourse. In: Sassen, C./ Benz. A./ Kühnlein, P. (eds.), *Constraints in Discourse*

Consten, M./ Knees, M./ Schwarz-Friesel, M.: forthc., The Function of Complex Anaphors in Texts. In: Schwarz-Friesel, M. / Consten, M./ Knees, M. (eds.), *Anaphors in Texts*

Cornish, F.: 1999, *Anaphora, Discourse, and Understanding: Evidence from English and French*. Oxford: Clarendon Press

Fraurud, K.: 1992, Situation Reference. What does ‚it' refer to? In: Fraurud, K., *Processing Noun Phrases in Natural Discourse*. PhD thesis. Depart. of Linguistics, Stockholm University

Hegarty, M.: 2003, *Type shifting of Entities in Discourse*. Presentation at the First International Workshop on Current Research in the Semantics-Pragmatics Interface, Michigan State University

Higginbotham, J.: 2000, On Events in Linguistic Semantics. In: Higginbotham, J. (ed). *Speaking of Events*. New York: Oxford Univ. Press. pp.49-80

Löbner, S.: 1985, Definites. In: *J. of Semantics 4*, pp.279-326

Maienborn, C.: 2003, *Die logische Form von Kopula-Sätzen*. Berlin: Akademie-Verlag

Schwarz, M.: 2001, Establishing Coherence in Text. Conceptual Continuity and Text-world Models. In: *Logos and Language*. Vol. II, No. 1 (2001). pp.15-24

# COMIC RELIEF FOR ANANKASTIC CONDITIONALS

TIM FERNANDO

Computer Science Department
Trinity College Dublin
Tim.Fernando@cs.tcd.ie

**Abstract**. Anankastic conditionals are analyzed in terms of events conceived as sequences of snapshots – roughly, comics. Quantification is applied not to worlds (sets of which are customarily identified with propositions) but to strings that record observations of actions. The account generalizes to other types of conditionals, sidestepping certain well-known problems that beset possible worlds treatments, such as logical omniscience and irrelevance. A refinement for anankastic conditionals is considered, incorporating action relations.

## 1. Introduction: over-generating means

Sentences such as (1) have recently attracted the attention of semanticists interested in the challenge they pose for the approach to modality described in Kratzer 1991.

(1) If you want to go to Harlem, you must take the A train. (Sæbø 2001)

(2) If you want to go to Harlem, you must take the A train to do that.

(3) To go to Harlem, you must take the A train.

Read as an *anankastic conditional*, (1) associates, according to von Fintel and Iatridou 2005, a goal with *must*, expressed in (2) and (3). Hence, (4) is deviant as an anankastic conditional; seeing the Apollo theatre is a "must" for visitors *already in* Harlem, *not* part of a means for getting to Harlem.

(4) If you want to go to Harlem, you must see the Apollo theatre.

"Being a means for" has, as noted in von Stechow et al. 2005, proved difficult for semantic accounts of anankastic conditionals based on possible worlds. Both von Stechow et al. 2005 and Huitink 2005 acknowledge problems with overgeneration; the former traces this to a lack of "understanding of the semantics of purpose clauses" while the latter argues that this is a job for pragmatics. Elaborating on what (2) adds to (1), Nissenbaum 2005 and von Fintel and Iatridou 2005 propose specific remedies, reviewed below. I claim that these works, taken together, suggest (albeit unwittingly perhaps) stepping from propositions down to actions, relations between which are employed in Balkanski 1992 and Di Eugenio and Webber 1996

to analyze "being a means for." The "relief" offered in the present paper draws on all the aforementioned accounts, plus Fernando 2004, where events (recording actions) are strung out as comics.

## 2. From propositions to actions

Let us, as usual, identify a proposition with the set of possible worlds where it is true, and agree that a modal base $m(w)$ of a world $w$ is a set of worlds accessible from $w$. A simple semantics for sentences such as (3) is (5).

(5) *to $p$, must $q$* is true in world $w$ relative to modal base $m(w)$ iff all worlds in $m(w)$ that belong to $p$ belong to $q$ — i.e. $m(w) \cap p \subseteq q$

If we set $m(w)$ to the intersection $\bigcap f(w)$ given by the conversational background function $f$ of Kratzer 1991, (5) becomes line (24b) in von Fintel and Iatridou 2005. (5) over-generates because it fails to link $q$ causally with $p$. Accordingly, von Fintel and Iatridou add to the right hand side of (5) the requirement that $q$ *be an essential part of a way of achieving* $p$, formalized relative to a world $w$ with modal base $m(w)$ as (6).

(6) for some set $P$ of propositions, $m(w) \cap q \cap \bigcap P \subseteq p$ but $m(w) \cap \bigcap P \not\subseteq p$

Unpacking (6), we can think of the set $P$ as a partial plan that lacks only $q$ to bring about $p$. But can we assume that time takes care of itself in (6)? Suppose, for the sake of the argument, that noone who goes to Harlem can leave Harlem — that is, $m(w)$ includes the proposition that everyone who goes to Harlem dies in Harlem. Would it be fair to say then that dying in Harlem is part of a way of getting to Harlem? As an anankastic conditional, (7) should be no truer than (4).

(7) If you want to go to Harlem, you must die in Harlem.

In general, if (6) holds for $q$, then it holds for $q \cap r$, for any $r$. Combined with (5), this may not be a problem if $r$ is not necessary given $p$. (Take seeing the Apollo theatre for $r$.) But who can rule out pesky conditions such as those supposed above for (7)? Or setting $q = p$ in (6), can we really construe (8) anankastically?

(8) If you want to go to Harlem, then you must go to Harlem.

And what about satisfying (6) with $q$ equal to $r \rightarrow p$, for $r$ in $P \cup f(w)$?

The brute fact is that a sentence such as (3) does not come so readily with propositional constituents $p$ and $q$. Its constituents $p$ and $q$ are arguably actions that combine temporally, not atemporally (as in $\cap$). While it is easy enough to express the performance of an action at a particular time as a proposition, it is not trivial to recover an action from a set of possible worlds. Nor is it clear that a set $q$ of worlds can meaningfully be part of a means for any set of worlds, unless we can associate an action with $q$.

But what is an action? One answer, not fully satisfactory but instructive nonetheless, is that an action is a program — something a programmer writes that is meant to be executed. A program $\pi$ for a commuter at train station X wishing to get to Harlem might consist of the 4 steps below.

$\pi$ = *walk to platform 2; board A train; ride A train; get off at 4th stop*

As with any description, the degree of detail in $\pi$ is bounded; we might introduce intermediate steps such as *wait at platform*, or perhaps concurrent actions such as *stay awake*. Exactly what instructions to state is a difficult matter that certainly calls for pragmatic reasoning (and more). This aside, there are two questions to ask about $\pi$ or indeed any program: whether we can carry it out (e.g. does the A train stop at platform 2?), and if we do, whether it would have the desired effect (i.e. is Harlem the A train's fourth stop from X?). For (1)-(3), what counts is not a program in paper (or in some mind), but a (complete) run of a program that transports us to Harlem.

## 3. Events as observations of actions

Let us fix a set $\Phi$ of temporal propositions, called *fluents* (McCarthy and Hayes 1969), such as *walk-to-platform2* saying that some agent (we leave implicit, for brevity) walks to platform 2. We hyphenate fluents to distinguish them from instructions that, for instance, appear in the 4-step program $\pi$ from the previous section. We can picture a run of $\pi$ as the string

$\hat{s}$ = $\boxed{\textit{walk-to-platform2}}\boxed{\textit{board-Atrain}}\boxed{\textit{ride-Atrain}}\boxed{\textit{get-off-at-4th-stop-from-X}}$

with substring $\boxed{\textit{board-Atrain}}\boxed{\textit{ride-Atrain}}$ describing an event of taking the A train.[1] In general, we construe a string $\alpha_1 \cdots \alpha_n \in \textit{Power}(\Phi)^+$ as an event of $n$ successive moments, with every fluent in $\alpha_i$ asserted to hold at the $i$th moment (for $1 \leq i \leq n$).[2] We extend inclusion $\supseteq$ between sets to strings in $\textit{Power}(\Phi)^+$, defining *subsumption* $\trianglerighteq$ to hold between strings of the same length when $\supseteq$ holds componentwise

$$\alpha_1 \cdots \alpha_n \trianglerighteq \beta_1 \cdots \beta_m \quad \text{iff} \quad n = m \textit{ and } \alpha_i \supseteq \beta_i \textit{ for } 1 \leq i \leq n\,.$$

For instance, $\hat{s} \trianglerighteq \square\boxed{\textit{board-Atrain}}\boxed{\textit{ride-Atrain}}\square$ (for $\hat{s}$ as above). Next, just as we form sets of possible worlds for propositions, we collect strings in languages

[1] Hence, $\hat{s}$ is a run also of the 3-step program *walk to platform 2; take A train; get off at 4th stop* . (No hyphens.)

[2] $\textit{Power}(\Phi)$ is the set of subsets of $\Phi$. To reinforce the intuition of a string as a comic-strip, boxes are used to enclose sets, understood as symbols from which strings are formed. Thus, the empty set is written as $\square$, when it is meant as a symbol, as opposed say, to the empty language $\emptyset$ containing no strings. Following the practice of regular expressions, we conflate a string $a$ with the singleton language $\{a\}$, lift concatenation to languages, and write Kleene star $^*$ for iteration (with $L^+ = LL^*$).

$L, L', \ldots$ over the alphabet *Power*$(\Phi)$ for event-types, and define $L$ to *subsume* $L'$ if every string in $L$ subsumes some string in $L'$

$$L \trianglerighteq L' \quad \text{iff} \quad (\forall s \in L)(\exists s' \in L')\ s \trianglerighteq s'$$

(Fernando 2004). We say $L$ *explicitly entails* $L'$ and write $L \vdash L'$ if $L \trianglerighteq \Box^* L' \Box^*$ (padding $L'$ with $\Box$'s to undo the requirement of equal length in $\trianglerighteq$). Identifying a string $s$ with the language $\{s\}$, we have $\hat{s} \vdash \boxed{\textit{board-Atrain}}\boxed{\textit{ride-Atrain}}$. Languages with any number of strings allow us to formulate an alternative to (5) simply as (9), where $\mathcal{W}_c(p)$ is a set of ways to *get to* $p$ from a starting point supplied by context $c$, while $\mathcal{L}(q)$ is the set of events of type $q$.

(9) *to p, must q* is true in context $c$ iff $\mathcal{W}_c(p) \vdash \mathcal{L}(q)$

The right hand side of (9) says every way to get to $p$ includes a $q$-event from a language $\mathcal{L}(q) \subseteq \textit{Power}(\Phi)^+$, which we may assume is given by linguistic knowledge (assembled, for instance, from a lexicon specifying the meanings of words in $q$). As for $\mathcal{W}_c(p)$, the necessity for going beyond $\mathcal{L}(p)$ is illustrated by (10), uttered in a context where speaker and addressee find themselves in a deserted island off Dublin with no food or drink or boat.

(10) To drink Guinness, you must swim to Dublin.

As no lexical semantics for drink can be expected to involve swimming, we must turn to context $c$ for a starting point for the actions in $\mathcal{W}_c(p)$, drawing on world knowledge along the way. But isn't (9) a bit ad hoc?

Far from it, I claim (9) generalizes to conditionals *if* $p$ *then* $q$, provided $\mathcal{W}_c$ is allowed to vary according to the kind of conditional (reading) involved. This is because it is natural to generate entailments by enriching explicit entailment $\vdash$ with maps $\mathcal{E}$ on languages $L$ as follows

$$L \vdash_{\mathcal{E}} L' \quad \text{iff} \quad \mathcal{E}(L) \vdash L' \,.$$

If, as one may expect, $\mathcal{E}$ elaborates on $L$ in that $\mathcal{E}(L) \vdash L$ then $\vdash$ is a subset of $\vdash_{\mathcal{E}}$. We might form $\mathcal{W}_c(p)$ as $\mathcal{E}_c(\mathcal{L}(p))$ for an $\mathcal{E}_c$ such that $\mathcal{E}_c(L) \vdash L$. For anankastic conditionals, the temporal span of $\mathcal{E}_c(L)$ may extend backward from $L$ to its past/left (as specified by $c$), but *not* forward, into the future/right.[3] This accounts for the oddness of (4) and (7) as anankastic conditionals; neither a visit to the Apollo theatre nor death in Harlem falls within the temporal span of $\mathcal{W}_c$(go to Harlem). Tailoring $\mathcal{E}$ according to the type of conditional of interest, we have a flexible handle on inference, with temporal matters strung out in full view. Since a string in $\mathcal{E}(L)$ falls short of a possible world, we may sidestep problems such as logical omniscience and irrelevance that plague possible worlds treatments. For instance, we may assume kissing Pedro Martinez (Nissenbaum 2005) does not appear in $\mathcal{W}_c$(go to Harlem).[4]

[3] That is, for *no* $s \in \mathcal{E}_c(L)$ do we have $s \trianglerighteq L\Box^+$. Hence, we can require $q$ to precede $p$ if we replace $\mathcal{L}(q)$ in (9) by $\mathcal{L}(q)\Box$, falsifying (8), where $p = q$.

[4] The obvious existential version of (9) is (9)′.

## 4. A refinement

Instead of sharpening a set of possible worlds where $p$ is achieved to some language $\mathcal{W}_c(p) \subseteq \mathit{Power}(\Phi)^+$, we might follow Nissenbaum 2005 and re-analyze $q$ as *$q$ with the goal to $p$*. This proposal, however, is flawed; as pointed out in von Fintel and Iatridou 2005 for the A train in (1), "it doesn't matter whether you take it with the goal of going to Harlem (as long as you get off at the right stop)" [p 19].

That said, we may nevertheless analyze $q$ alongside $p$, replacing $\mathcal{L}(q)$ in (9) by a set $\mathcal{R}(q,p)$ of runs of $q$ that lead to $p$ in that

$$\mathcal{R}(q,p) \;\trianglerighteq\; \mathcal{L}(q)\Box^* \quad \text{and} \quad \mathcal{R}(q,p) \;\trianglerighteq\; \Box^*\mathcal{L}(p)\ .$$

(11) *to p, must q* is true in context $c$ iff $\mathcal{W}_c(p) \;\vdash\; \mathcal{R}(q,p)$

Some support for changing $\mathcal{L}(q)$ in (9) to $\mathcal{R}(q,p)$ in (11) is to be found in the observation from Di Eugenio and Webber 1996 that in the phrase *$q$ to $p$*, the goal $p$ constrains $q$ (presumably under minimal rationality assumptions). These constraints can be imposed through action relations linking $q$ to $p$, which Balkanski 1992 argues are (among other things) irreflexive, marking *$p$ to $p$* as odd. Thus, if (8) is indeed deviant as an anankastic conditional, we have a ready explanation from (11) and the constraint $\mathcal{R}(p,p) = \emptyset$.[5] In Di Eugenio and Webber 1996, action relations bind actions to form *plan graphs*, representing intentions. The aforementioned problem with Nissenbaum's proposal suggests we should be careful about reducing these intentions to claims entirely about mental states. Both (9) and (11) focus on what can be observed from the outside as strings over the alphabet $\mathit{Power}(\Phi)$. But action relations may well call for greater boldness in speculating about what lies inside the black box, so as to understand how black boxes might interact to produce, for example, $\mathcal{R}(q,p)$.

## 5. Conclusion: from worlds to strings

Proposals (9) and (11) above reduce the semantic over-generation of previous accounts based on possible worlds that are arguably too crude to capture what "being a means for" means. The step from worlds down to strings recording observations of a computational/cognitive mechanism is an attempt to stake out some middle ground between what Jackendoff 1996 calls E[xternalized]-semantics and I[nternalized]-semantics — between on the one hand, propositions and truth (against some external world), and on the other hand, actions and computation/cognition.

(9)′ *to p, may q* is true in context $c$ iff $(\exists s \in \mathcal{W}_c(p))\ s \vdash\ \mathcal{L}(q)$

To interpret *ought* relative to a preference relation $\leq_c$ on strings, let $(Q^{\leq_c} s \in L)\ A(s)$ abbreviate $(\forall s \in L)\ (\exists s' \leq_c^L s)(\forall s'' \leq_c^L s')\ A(s'')$, where $s \leq_c^L s'$ abbreviates $s \leq_c s'$ *and* $s \in L$.

(9)″ *to p, ought q* is true in context $c$ iff $(Q^{\leq_c} s \in \mathcal{W}_c(p))\ s \vdash\ \mathcal{L}(q)$

[5] See also footnote 3 for an explanation covering the case where the action relation is enablement.

Pragmatic input to $\mathcal{W}_c(p)$ and, in the case of (11), to $\mathcal{R}(q,p)$ breaks the traditional semantics-pragmatics pipeline that places semantic matters of truth strictly ahead of pragmatic considerations of use.[6]

## Bibliography


Balkanski, C.: 1992, Action relations in rationale clauses and means clauses, in *Proc. COLING-92*, pp 267–273

Di Eugenio, B. and Webber, B.: 1996, Pragmatic overloading in natural language instructions, *International Journal of Expert Systems* 9(1), 53–84

Fernando, T.: 2004, A finite-state approach to events in natural language semantics, *Journal of Logic and Computation* 14(1), 79–92

von Fintel, K. and Iatridou, S.: September 2005, *What to do if you want to go to Harlem: anankastic conditionals and related matters*, Draft for Rutgers Semantics Workshop

Huitink, J.: 2005, Analyzing anankastic conditionals and sufficiency modals, in *Proc. 13th conference of the Student Organization of Linguistics in Europe (SOLE)*

Jackendoff, R.: 1996, Semantics and cognition, in S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, pp 539–559, Blackwell, Oxford

Kratzer, A.: 1991, Modality, in A. von Stechow and D. Wunderlich (eds.), *Semantics*, pp 639–650, Walter de Gruyter, Berlin

McCarthy, J. and Hayes, P.: 1969, Some philosophical problems from the standpoint of artificial intelligence, in M. Meltzer and D. Michie (eds.), *Machine Intelligence 4*, pp 463–502, Edinburgh University Press

Nissenbaum, J.: 2005, Kissing Pedro Martinez: (existential) anankastic conditionals and rationale clauses, in *Proc. Semantics and Linguistic Theory XV*, Cornell Linguistics Circle Publications

Sæbø, K. J.: 2001, Necessary conditions in natural language, in C. Féry and W. Sternefeld (eds.), *Audiatur Vox Sapientiae: A Festschrift for Arnim von Stechow*, pp 427–449, Akademie Verlag, Berlin

von Stechow, A., Krasikova, S., and Penka, D.: April 2005, *Anankastic conditionals*, Handout for a talk at Cornell


[6] **Post Script**: (12) and (13) pose problems for (9)/(11), brought to my attention by Magdalena Schwager.

(12) If you want to go to Harlem, you must not take the B train.

(13) If you want to go to Harlem, you must stay awake.

For (12) and (13), it is unfortunate that the temporal force of $q$ in (9)/(11) should be made existential by the step from $\trianglerighteq$ to $\vdash$. While I cannot claim to have a fully worked out proposal at the moment, I am, for the case of (9), inclined to put into $\mathcal{L}(q)$ a fluent that, for (12) and (13), has effects beyond the box that encloses it — as is the case for inertia and fluents with temporal operators (discussed in my SALT 2004 and FG/MOL 2005 papers, available in `www.cs.tcd.ie/Tim.Fernando`). The $q$'s in (12) and (13) are, after all, stative. Whether this approach can be pulled off cleanly, I regret I cannot say.

# ACHIEVING EXPRESSIVE COMPLETENESS AND COMPUTATIONAL EFFICIENCY FOR UNDERSPECIFIED SCOPE REPRESENTATIONS

CHRIS FOX AND SHALOM LAPPIN

Department of Computer Science
University of Essex
foxcj@essex.ac.uk

Department of Philosophy
King's College London
shalom.lappin@kcl.ac.uk

The tension between expressive power and computational tractability poses an acute problem for theories of underspecified semantic representation. In previous work we have presented an account of underspecified scope representations within Property Theory with Curry Typing (PTCT), an intensional first-order theory for natural language semantics. Here we show how filters applied to the underspecified-scope terms of PTCT permit both expressive completeness and the reduction of computational complexity in a significant class of non-worst case scenarios.

## 1. Introduction

In Fox and Lappin (2005a) we propose Property Theory with Curry Typing (PTCT) as a formal framework for the semantics of natural language. PTCT allows fine-grained distinctions of meaning without recourse to modal notions like (im)possible worlds. It also supports a unified dynamic treatment of pronominal anaphora and VP ellipsis, as well as related phenomena such as gapping and pseudo-gapping.

PTCT consists of three sublanguage components. The first component encodes a property theory within a language of terms (an untyped $\lambda$-calculus). The second adds dynamic Curry typing (Curry and Feys, 1958) to provide a system for expressing type judgements for terms. The third uses a first-order logic to specify the truth-conditions of the propositional subpart of the term language. Our semantic representation language is first-order in character, rather than higher-order. We achieve the sort of expressive power previously limited to higher-order theories within a formally more constrained system. This provides an effective procedure for modelling inference in natural language.

In Fox and Lappin (2005a,b) product types are used to generate underspecified semantic representations within PTCT, the representation language, rather than through meta-language devices, which are invoked in most current treatments of underspecification (Reyle, 1993; Bos, 1995; Blackburn and Bos, 2005; Copestake et al., 1997). The expressive power of the language permits the formulation of filters on scope readings that cannot be captured in other theories of underspecification which rely on special purpose extra-linguistic operations and a weak system for constraint specification.

These filters on underspecified scope terms can solve the problem of expressive incompleteness that Ebert (2005) raises for other theories of underspecification. They can also be used to reduce the complexity involved in computing the set of possible scope readings that an underspecified term generates.

## 2. PTCT

PTCT is a first-order sysem in which types and propositions are terms over which we can quantify. This allows rich expressiveness whilst restricting the system to first order resources (Fox and Lappin, 2005a, Chapter 9).

The language of terms is the untyped $\lambda$-calculus, enriched with logical constants. It is used to *represent* the interpretations of natural language expressions. It has no internal logic, but when we add a proof theory, the simple language of types together with the language of terms can be combined to produce a Curry-typed $\lambda$-calculus.

The syntactic rules of PTCT are flexible. They allow the generation of syntactic expressions that have no intuitively meaningful interpretation. This does not undermine the system. The rules give a minimal characterisation of the syntax while our proof theory and our model theory characterise the proper subset of well-formed PTCT terms that constitute meaningful expressions.

In the first-order language of wffs we formulate type judgements for terms, and truth conditions for those terms judged to be in Prop.

## 3. Underspecified Representations in PTCT

Generalised quantifiers (GQs) represent noun phrases. We follow Keenan (1992) and van Eijck (2003) in taking a GQ to be an arity reduction operator that applies to a relation $r$ to yield either a proposition or a relation $r'$ that is produced by effectively saturating one of $r$'s argument with the GQ.

We specify a family of functions $perms\_scope_k$ (where $k > 1$) that generate all $k!$ indexed permutation products of a $k$-ary indexed product term $\langle t_1, \ldots, t_k \rangle$ as part of the procedure for generating the set of possible scope readings of a sentence.

For our treatment of underspecification, $perms\_scope_k$ needs to take a $k$-ary product of scope taking elements (by default, in the order in which they appear in the surface syntax) and a $k$-ary relation representing the core proposition as its arguments. The scope taking elements and the core representation can be combined into a single product, e.g. as a pair consisting of the $k$-tuples of quantifiers as its first element and the core relation as its second. The permutation function $perms\_scope_k$ produces the $k!$-ary product of scoped readings. When a $k$-tuple of quantifiers is permuted, the $\lambda$-operators that bind the quantified argument positions in the core relation are effectively permuted in the same order as the quantifiers in the $k$-tuple. This correspondence is necessary to preserve the connection between each GQ and its argument position in the core relation across scope permutations.

A scope reading is generated by applying the elements of the $k$-tuple of quantifiers in sequence to the core relation, reducing its arity with each such operation until a proposition results. The $i$th scope reading is identified by projecting the $i$th element of the indexed product of propositions that is the output of our $perms\_scope_k$ function. Therefore, the PTCT term consisting of the application of $perms\_scope_k$ to an input pair of a $k$-tuple of GQs and a core relation provides an underspecified representation of the sentence corresponding to this term.

## 4. Filters and Expressive Completeness

Scope constraints can be formulated as filters on the $k!$-tuple of permutations of the form $\langle\langle Qtuple_1, Rel_1\rangle, \ldots, \langle Qtuple_{k!}, Rel_{k!}\rangle\rangle$ that $perms\_scope_k$ generates for an argument pair $\langle Qtuple_1, Rel_1\rangle$. Each such filter is a Boolean property function that imposes a condition on the elements of the $k!$-tuple.

Underspecified representations can be disambiguated by information acquired through subsequent discourse. So, for example, resolving anaphoric expressions like pronouns and definite descriptions in sentences following a statement that exhibits scope ambiguity may eliminate certain readings of the antecedent.

(1) A: Every student wrote a program for some professor.
(2) B: Yes, I know the professor. She taught the Haskell course.
(3) C: I saw the programs, and they were all list-sorting procedures.

Identifying "*some professor*" in (1) as the antecedent for "*the professor*" and "*she*" in (2) gives "*some professor*" scope over "*every student*" in (1). Interpreting "*a program*" in (1) as the antecedent for "*the programs*" and "*they*" in (3) causes "*a program*" to have narrow scope relative to "*every student*" in (1). Therefore, taken conjointly (2) and (3) forces on (1) the fully resolved scope order

$$\langle \text{“}\textit{some professor}\text{”}, \text{“}\textit{every student}\text{”}, \text{“}\textit{a program}\text{”} \rangle$$

Assume that "*every student*" $= Q_1$, "*a program*" $= Q_2$, and "*some professor*" $= Q_3$. We can formulate the filters contributed by (2) and (3) as (4) and (5), respectively (where $GQ$ in $\hat{=}_{GQ}$ abbreviates the appropriate type of $Q_i$). In these filters we take $\langle Quants, Rel \rangle$ to be a variable ranging over pairs in which $Quants$ is a $k$-tuple and $Rel$ is a $k$-ary relation. As the $k$-tuples are indexed, there is a one-to-one correspondence between the elements of a $k$-tuple and their respective indices. Let $tuple\_elem(i, Quants) = Q_i$ if $Q_i$ is the $i$th member of $Quants$, and the distinguished term $\omega$ otherwise.

(4) $\lambda\langle Quants, Rel\rangle[\hat{\forall} i \epsilon \mathsf{Num} \hat{\forall} j \epsilon \mathsf{Num}((tuple\_elem(i, Quants) \hat{=}_{GQ} Q_3 \hat{\wedge}$
$tuple\_elem(j, Quants) \hat{=}_{GQ} Q_1) \hat{\rightarrow}$
$i \hat{<} j)]$

(5) $\lambda\langle Quants, Rel\rangle[\hat{\forall} i \epsilon \mathsf{Num} \hat{\forall} j \epsilon \mathsf{Num}((tuple\_elem(i, Quants) \hat{=}_{GQ} Q_1 \hat{\wedge}$
$tuple\_elem(j, Quants) \hat{=}_{GQ} Q_2) \hat{\rightarrow}$
$i \hat{<} j)]$

We specify the function $filter\_tuple(\langle F, T\rangle)$ which maps a pair consisting of a $j$-tuple $F$ of filters and a $k$-tuple $T$ to a $k'$-tuple (possibly the empty tuple) of all the elements of $T$ that satisfy each filter in $F$. We construct a PTCT term of the form (6) to represent the $k'$-tuple obtained by applying the elements of $F$ to the $k!$-tuple that is the value of $perms\_scope_k(\langle Quants_k, Rel\rangle)$.

(6) $filter\_tuple(\langle F, perms\_scope_k(\langle Quants_k, Rel\rangle)\rangle)$

Ebert (2005) shows that most current theories of underspecification are expressively incomplete to the extent that they cannot identify the proper subset of possible scope readings specified by Boolean operations other than conjunction, and in particular by negation. He cites the following example to illustrate the problem.

(7) Every market manager showed five sales representatives a sample.

Ebert stipulates that, in his example, real world knowledge allows all scope permutations except the one corresponding to $\langle \exists, 5, \forall \rangle$, where *a sample* takes wide scope, *five sales representatives* intermediary position, and *every market manager* narrow scope. He demonstrates that storage (Cooper, 1983; Pereira, 1990), hole semantics (Bos, 1995; Blackburn and Bos, 2005), Minimal Recursion Semantics (Copestake et al., 1997), and Normal Dominance Conditions (Koller et al., 2003) cannot formulate underspecified representations that express the set containing only the five remaining scope readings.

By contrast it is straightforward to formulate a filter in PTCT that rules out the problematic scope sequence in Ebert's case while permitting the five other readings.

(8) $\lambda\langle Quants, Rel\rangle[\hat{\forall} i \epsilon \mathsf{Num} \hat{\forall} j \epsilon \mathsf{Num} \hat{\forall} k \epsilon \mathsf{Num}((tuple\_elem(i, Quants) \hat{=}_{GQ} Q_{\exists}$
$\hat{\wedge}\ tuple\_elem(j, Quants) \hat{=}_{GQ} Q_5$
$\hat{\wedge}\ tuple\_elem(k, Quants) \hat{=}_{GQ} Q_{\forall}) \hat{\rightarrow} \hat{\sim}(i \hat{<} j \hat{\wedge} j \hat{<} k))]$

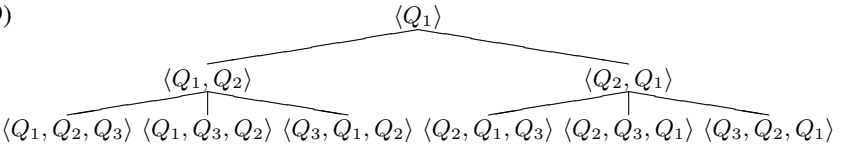
PTCT is, in principle, able to achieve expressive completeness in Ebert's (2005) sense.

## 5. Efficient Computation of Possible Scope Readings

At first glance it might seem that it is, in general, necessary to generate the full $k!$-tuple that is the value of $perms\_scope_k(\langle Quants_k, Rel\rangle)$ before applying the filters of $F$ to the elements of this $k!$-tuple in order to compute the value of (6). Fortunately, this is not the case.

In Fox and Lappin (2005c) we present a tree construction algorithm for generating all possible permutations of a $k$-tuple. If this algorithm takes as its input the triple $\langle Q_1, Q_2, Q_3\rangle$, then it generates the following tree.

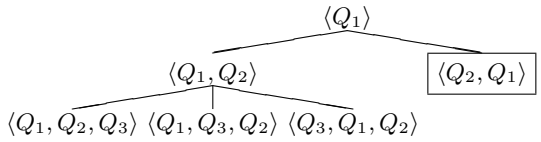
(9)




Filters can apply as constraints to nodes in the tree as the algorithm produces them. If a node violates a filter, then it is deleted, and the subtree that it dominates is not generated. In this way filters can reduce the size of the tree, and so limit the search space of possible scope readings that are explored for underspecified-scope terms $perms\_scope_k(\langle Quants_k, Rel\rangle)$ to a proper subset of the elements of the $k!$-tuple that is its value.

So, for example, the filter $Q_1 < Q_2$ prunes the tree in (9) to give the one in (10).

(10)




Identifying the size of a tree with the number of its nodes, we can compute the size of a tree $T$, $|T|$, through the formula

(11) $|T| = \Sigma i!$, where $i$ is the index of the $i$th element of the initial $k$-tuple which the algorithm takes as its input.

Therefore, the size of the tree in (9) is $1! + 2! + 3! = 9$. The size of the tree in (10) is 6, which is a reduction of 30%.

The size of a subtree $ST$ dominated by a node $n$ at level $i$, but not including $n$, is given by the formula

(12) $|ST| = \Pi j\ (i < j \leq k) + \Sigma j'\ (i < j' < k)$.

Consider the quadruple $\langle Q_1, Q_2, Q_3, Q_4\rangle$. The tree algorithm produces an indexed $k$!-tuple of 24 $k$-tuples as the leaves of a tree $T_4$ with 4 levels and 33 nodes. If a filter like $Q_1 < Q_2$ applies at level 2, the first branching node of $T_4$, it prunes the right-half of $T_4$ under $\langle Q_2, Q_1\rangle$, and so it eliminates a subtree of 15 nodes, reducing $T_4$ by 15/33 = 45.4%. The remaining left side of $T_4$ has the three nodes $\langle Q_1, Q_2, Q_3\rangle$, $\langle Q_1, Q_3, Q_2\rangle$, $\langle Q_3, Q_1, Q_2\rangle$ at level 3. If the filter $Q_2 < Q_3$ applies at this level, the 8 leaf nodes under $\langle Q_1, Q_3, Q_2\rangle$ and $\langle Q_3, Q_1, Q_2\rangle$ are pruned. Therefore, the conjunction of the filters $Q_1 < Q_2$ and $Q_2 < Q_3$ reduces $T_4$ by 15 + 8 = 23 nodes, which is (approximately) 70% of the full tree.

It is not difficult to construct a plausible case in which the interpretation of a sentence containing four quantified NPs is disambiguated by a conjunction of two filters of this kind through anaphora resolution in subsequent discourse, as in **A:** "It's amazing. A critic recently reviewed two plays for every newspaper in a major city." **B:** "Yes, he published the same reviews of *A Midsummer Night's Dream* and *New-Found-Land* in every major paper in New York last week."

Clearly, the earlier in the tree construction process (the higher up in the tree) that a filter applies, the greater the reduction in search space of possible scope readings that it achieves. It is also possible to optimise the interaction of filters and the tree construction algorithm by specifying a procedure that reorders the elements of the input $k$-tuple to permit the filters to apply at the earliest point in the generation of the tree. So, for example, if the algorithm takes as its input the triple $\langle Q_1, Q_2, Q_3\rangle$ and one of the filters that apply to this triple is $Q_2 < Q_3$, then the reordering operation will map the triple into $\langle Q_2, Q_3, Q_1\rangle$. We will leave the formulation of this operation for future work.

Ebert (2005) proves a theorem that entails that if a theory is expressively complete, then it will, in the worst case, produce a combinatorial explosion equivalent to generating all $k!$ scope readings for a sentence. This result holds for PTCT in the limit case, where no filters have been applied to a $perms\_scope_k(\langle Quants_k, Rel\rangle)$ term, or they do not operate early enough in the tree construction algorithm to restrict the scope permutation tree. However, as we have seen, there is a large class of cases in which filters significantly reduce the search space through tree pruning, and so they offer a mechanism for rendering scope disambiguation computationally efficient.

## 6. Conclusion

We have formulated constraints on scope readings as filters on the $k$!-tuples that $perms\_scope_k$ produces. These filters are PTCT terms which encode Boolean conditions and quantification over the integers of indexed $k$-tuples. In principle, they permit PTCT to achieve expressive completeness in the sense of Ebert (2005).

We have also invoked a tree generation algorithm to characterise (the permutation part of) the computable function that $perms\_scope_k$ denotes. When filters are applied as constraints on nodes in the tree that the algorithm generates, they can significantly reduce the search space of possible scope readings given by an underspecified representation.

Underspecified representations, the projection of a particular scope interpretation, and constraints on possible scope readings are all specified by appropriately typed $\lambda$-terms within the semantic representation language, PTCT, rather than through operations on schematic metalinguistic objects. Our proposed treatment of underspecified representations within PTCT achieves both significant expressive power and efficient computation of possible scope interpretations.

**Acknowledgements**

We are grateful to Christian Ebert for helpful comments on an earlier draft of this paper. His recent Ph.D. dissertation has stimulated and influenced much of the work we report here. We would also like to thank Dick Crouch and Ron Kaplan for useful discussion of the complexity issues we address. The second author's research was supported by research grant number RES–000–23–0065 from the Economic and Social Research Council of the United Kingdom.

## References

Blackburn, P. and Bos, J.: 2005, *Representation and Inference for Natural Language*, CSLI, Stanford

Bos, J.: 1995, Predicate logic unplugged, in *Proceedings of the Tenth Amsterdam Colloquium*, Amsterdam, Holland

Cooper, R.: 1983, *Quantification and Syntactic Theory*, Synthese Language Library, D. Reidel, Dordrecht

Copestake, A., Flickinger, D., and Sag, I. A.: 1997, *Minimal Recursion Semantics*, Tech. rep., Stanford University, Stanford, CA, unpublished ms.

Curry, H. B. and Feys, R.: 1958, *Combinatory Logic*, vol. 1 of *Studies in Logic*, North Holland

Ebert, C.: 2005, *Formal Investigation of Underspecified Representations*, Ph.D. thesis, Department of Computer Science, King's College London, unpublished

van Eijck, J.: 2003, *Computational Semantics and Type Theory*, unpublished ms., CWI, Amsterdam

Fox, C. and Lappin, S.: 2005a, *Formal Foundations of Intensional Semantics*, Blackwell, Oxford

Fox, C. and Lappin, S.: 2005b, Underspecified interpretations in a Curry-typed representation language, *The Journal of Logic and Computation* 15, 131–143

Fox, C. and Lappin, S.: 2005c, *Expressive Completeness and Computational Efficiency for Underspecified Representations*, unpublished ms., University of Essex and King's College London

Keenan, E.: 1992, Beyond the Fregean boundary, *Linguistics and Philosophy* 15, 199–221

Koller, A., Niehren, J., and Thater, S.: 2003, Bridging the gap between underspecified formalisms: Hole semantics as dominance constraints, in *Proceedings of 11th EACL*, Budapest

Pereira, F.: 1990, Categorial semantics and scoping, *Computational Linguistics* 16, 1–10

Reyle, U.: 1993, Dealing with ambiguities by underspecification: Construction, representation and deduction, *Journal of Semantics* 10, 123–179

# HOW AND HOW NOT TO EMPLOY DISCOURSE RELATIONS TO ACCOUNT FOR PSEUDO-IMPERATIVES

MICHAEL FRANKE

University of Amsterdam, ILLC
M.Franke@uva.nl

## 1. Introduction

Pseudo-imperatives are compound sentences where an imperative clause is conjoined or disjoined with a declarative clause (1).

(1) a. Do $X$, and $Y$ will happen/be the case/be done.
b. Do $X$, or $Y$ will happen/be the case/be done.

These constructions arouse interest because of an intriguing asymmetry in meaning and felicity. Although a pseudo-imperative with conjunction (1a) can be interpreted as asking for performance or omission of action $X$ depending on the desirability of the variable $Y$, pseudo-imperatives with disjunction (1b) require an instantiation of $Y$ for pragmatic felicity which is undesirable for the addressee, and cannot be used to request forbearance from $X$. It therefore needs to be asked:

(Q1) Why can (1a) mean "Don't do $X$!" given appropriate instantiation of $Y$?
(Q2) Why can (1b) *not* mean "Don't do $X$!" however $Y$ is instantiated?
(Q3) Why are instantiations of $Y$ in (1b) pragmatically infelicitous in case $Y$ cannot be interpreted as denoting an undesirable state of affairs for the addressee?

To resolve these issues, this paper argues for a uniform conditional-like treatment of pseudo-imperatives and suggests an account of the pragmatic asymmetry in terms of discourse relations.

## 2. Preliminaries

I will restrict myself to a discussion of pseudo-imperatives which contain *simple imperative forms*. A simple imperative form is what in English could be a bare VP[1], with the exception of the copula. This is meant to exclude imperative forms with do-support, special intonation or otherwise modified with lexical additions such as *please*, *will you*, *damn it* etc. This restriction excludes from consideration instantiations of (1a) such as (2) which behave like (1b) in that they cannot be interpreted to mean "Don't do $X$!" and are infelicitous under not-desirable instantiations of $Y$.

[1] Although English imperative forms cannot be distinguished from bare VPs, I assume here that what occurs in pseudo-imperatives are not bare VPs, but imperative forms, (i) because the English copula occurs in its unambiguous imperative form in pseudo-imperatives, and (ii) because of the fact that languages with unambiguously marked imperative forms in which pseudo-imperatives occur, e.g. German or Dutch, feature imperative forms and not bare VPs in pseudo-imperatives.

(2) ? Shut up, please, and I'll beat you.

Let upper-case $A$ denote simple imperative forms and let $P$ denote declarative sentences. Lower-case letters refer to semantic denotations. Following Mastop 2005, I will assume that imperative forms semantically denote actions. Lower-case $a$ is then used to refer to an action. Lower-case $p$ refers to a proposition. Pseudo-imperatives with conjunction (1a) are henceforth called IaDs, which is short for 'imperative and declarative', and schematically represented as "$A$ and $P$". Similarly, IoDs (1b) are represented as "$A$ or $P$".

Imperative forms are interpreted in a variety of ways and quite heterogeneous speech-acts are associated with utterances of imperative forms. I will in the following make a simplified distinction between institutional and descriptive uses of imperative forms. In institutional uses the imperative form is used to establish institutional facts. Institutional use of (3a), for instance, affects what the addressee may or may not do: he must not speak. I will write $!a$ for institutional uses of $A$ and I will say that an imperative $A$ *is given* iff it is institutionally used.

(3) a. Shut up, damn it!
b. If you don't want to lose your job, never let this happen again.
c. How can I win Jane's heart? - Dance the duck-dance, for instance.

In descriptive uses of imperative forms like (3b,c) institutional facts remain unaffected. The speaker merely describes a state of affairs: performance or omission of an action has or might have certain consequences. Such statements may certainly still be inducements or deterrents, as (3b) clearly shows. But this does not mean that a particular speech-act of inducing or deterring is associated with the imperative form alone. Rather, the contribution of a descriptively used imperative form $A$ is its semantic denotation $a$ which is asserted to stand in a particular consequence relation to a proposition $p$, and this assertion may in turn serve as an argument to influence the addressee's choice of action.

I propose to treat such consequence relations which hold between an action and a proposition as content-level discourse relations and to investigate which relations are featured in the interpretation of pseudo-imperatives. To this end, I will shortly introduce relevant discourse relations for which I will provide a very simplistic formal definition on the basis of a context-model: Let $W$ be a non-empty set of future courses of events over a fixed finite amount of time that are considered normal or salient in context. For a proposition $p$, let's write $w \models p$ if $p$ is an outcome of $w$. Let's write $a \in w$ if action $a$ is performed in $w$ and, if $a \in w$, let $w/a$ be the future course of events which is as much as $w$ except that $a$ is not performed. Let's write $\neg a$ for forbearance from $a$, so that $\neg a \in w$ iff $a \notin w$. Forbearance from an action is considered an action. Then define:

$$
\begin{array}{lll}
[a]p & \text{iff} & \forall w \in W\ (a \in w \rightarrow w \models p) \\
\textit{cause}(a,p) & \text{iff} & \exists w \in W\ (a \in w\ \wedge\ w \models p\ \wedge\ w/a \not\models p) \\
\textit{result}(a,p) & \text{iff} & [a]p \wedge \textit{cause}(a,p) \\
\textit{require}(a,p) & \text{iff} & [\neg a]p \wedge \textit{cause}(\neg a,p) \wedge \textit{undesirable}(p)
\end{array}
$$

As $W$ is meant to model normal courses of events, consequence relations are defeasible. The relation *result*$(a, p)$ expresses that the action $a$ brings about the state of affairs $p$. Reference to causality is needed to distinguish results from logical consequences, yet causality is to be kept separate from necessity. The relation *require*$(a, p)$ expresses that $a$ is necessary because otherwise $p$ holds. Unlike *result*, the latter relation contains a reference to the undesirability of $p$. The intuition here is that '$a$ is necessary, because otherwise $p$' is a natural notion only if $p$ is something to be avoided.

## 3. The Conjunctive Case

For an IaD "$A$ and $P$" I will argue that we do not have to assume that an imperative is given but that $A$ is only descriptively used and that a conditional interpretation $[a]p$ enriched to $result(a, p)$ is sufficient to account for intuitions.

It has been noted that IaDs, come in three flavors. Clark 1993 has termed these positive (4a), negative (4b) and neutral readings (4c).

(4) a. Come to my BBQ and you'll finally meet the minus girls.
b. Say one more word about minus girls and I'll throw you into the canal.
c. Swim in the canals regularly and you will live 3.14159 years longer.

In positive readings of an IaD "$A$ and $P$" the intuitive impact is to have the addressee perform the action $a$, while in negative readings the overall impact is to have the addressee omit the action $a$. In neutral readings no urge towards performance or omission is felt.

It is commonplace to accept that IaDs have some sort of conditional reading. Opinions differ, however, about the meaning contribution of the imperative form. For instance, van der Auwera 1986 maintains that the imperative in an IaD is always given, so that "$A$ and $P$" is to be analyzed as a speech-act conjunction, write $\&$, to yield $!a \;\&\; [a]p$. In negative readings of IaDs, van der Auwera then argues, if $p$ is undesirable for the hearer, we infer that $[a]p$ is a sufficiently strong argument to realize that not $!a$ is meant, but $!\neg a$. For neutral readings of IaDs, van der Auwera holds that "some [IaDs] seem to be primarily imperative, while others seem primarily conditional." (p. 209).

Lascarides and Asher 2004, on the other hand, hold that only in positive IaDs an imperative is given. They argue that the non-veridical discourse relation *Def-Consequence* holds between conjuncts in "$A$ and $P$", which corresponds to $[a]p$ in our terms. A further default interpretation rule then yields the veridical 'meta-talk'-relation *Explanation*$^*$ just in case $p$ is hearer-desirable.[2] Veridicality of $Explanation^*$ then ensures that the imperative $!a$ is given and that $[a]p$ explains why it was.

Against Lascarides and Asher, I argue that appeal to *Explanation*$^*$ is implausible, but also not necessary. It is implausible, because in positive IaDs the speaker is

[2]Recall that a meta-relation $R^*$ holds between two discourse units just in case the content of the one stands in relation $R$ to the fact that the speech act associated with the utterance of the other was performed.

felt to persuade the hearer into performance of an action $a$, but that surely does not mean that the incentive which the speaker gives to have the hearer perform $a$ is the speaker's own reason for the inducement, or even an imperative $!a$. Appeal to *Explanation*$^*$ is also not necessary, because we do not need to assume that an imperative is given in positive readings of IaDs. The only thing that matters for a positive interpretation of "$A$ and $P$" is that it influences the hearer to perform $a$. A statement $[a]p$ will influence the hearer to perform $a$ if $p$ can be construed, not only as a desirable logical consequence, but as a desirable result of $a$. Thus, "$A$ and $P$" will be an argument to perform $a$ just in case the content-level discourse relations $result(a, p)$ and $desirable(p)$ can be inferred. On top of that it is natural to assume that the speaker used "$A$ and $P$" as an inducement, which could be captured in the further ascription of an intentional-level discourse relation *Motivation*$(a, p)$ (Mann and Thompson 1987). However, two things should not be confused here. Surely, under normal circumstances, we infer that the speaker's intention behind an utterance of "$A$ and $P$" with hearer-desirable $p$ is to persuade the hearer to perform $a$. But that does not mean that we have to associate illocutionary force with $A$. In line with the terminological conventions introduced above, in positive IaDs no imperative is given. It suffices to treat the imperative form as descriptively used.

Negative readings can be accounted for in a similar fashion. We get a negative reading for "$A$ and $P$" just in case the proposition $p$ is interpretable as something to be avoided by forbearance from $a$. This is again sufficiently captured by the content-level relations *result*$(a, p)$ and *undesirable*$(p)$. The fact that for undesirable $p$ an IaD is felt to be an urge to forbear from $a$ neither justifies nor necessitates the association with institutional force $!\neg a$. In line with the above, we should rather ascribe an intentional-level discourse relation *Discourage*$(a, p)$, accordingly defined.

Finally, neutral readings of "$A$ and $P$" are exactly those where $p$ is neither construable as desirable, nor undesirable.[3] Needless to say, that in this case we cannot say that an imperative is given. The discourse relation *result*$(a, p)$ alone suffices to reflect the intuitive meaning.

In sum, I meant to propose that IaDs may be treated uniformly as result statements. The speaker's intention to influence the hearer's choice of action does not require that we associate illocutionary force with the featured imperative forms.

## 4. The Disjunctive Case

For an IoD "$A$ or $P$" I will argue that a mere conditional interpretation $[\neg a]p$ may be maintained for uniformity of analysis, if we assume that IoDs differ from conditional statements in that they require their second disjunct to relate to the topic addressed in the first, so that we find a plausible answer to questions (Q2) and (Q3) in the intuition that a conditional $[\neg a]p$ relates to a topic $a$ only in case $p$ is negatively connoted.

[3] Given the necessary space, the neutral case clearly deserved further attention. Neutral IaDs are often generic statements, rather than referring to an immediate action of the addressee.

Although an IoD "$A$ or $P$" is certainly associated with a conditional statement $[\neg a]p$, it seems implausible to assume that IoDs are merely such conditional statements, in light of the acceptability difference of a contrast pair like (5):

(5) a. ? Invite Jason, or we'll have more beer for ourselves.
b. If you do not invite Jason, we'll have more beer for ourselves.

The difference between IoDs and conditional statements may be sought in a speech-act conjunction analysis in parallel to other conditional uses of disjunction (6).

(6) a. It's a good idea to invite Jason, or we'll be in trouble finishing all that beer.
b. I will invite Jason, or we'll be in trouble finishing all that beer.
c. Jane hopes that Jason is coming, or she'll have to dance (the d.d.) alone.

The supposed parallel would then suggest to analyze an IoD "$A$ or $P$" as a speech-act conjunction $!a \;\&\; [\neg a]p$. The semantic role of disjunction in such speech-act conjunction readings might be characterized as an epistemic context-splitter modulo topic (cf. Schwager 2004). The first disjunct is about an epistemically uncertain state of affairs. It addresses, as a topic, the performance of an action in the future. Disjunction then supplies the negated topic worlds for contextual restriction of the modal in the second disjunct.

If we treat IoDs as a speech-act conjunction, we can account for the problems addressed in (Q2) and (Q3). The answer to question (Q2) is straightforward under the assumption that an imperative form $A$ is never associated with an imperative $!\neg a$. An answer to question (Q3) could then be that, at least for undesirable $p$, the conjoined speech acts are associated with incongruous intentions: whereas $!a$ is an inducement to perform $a$, the statement $[\neg a]p$ is a deterrent.

However there are at least two complications. For one thing, at least some IoDs are not associated intuitively with institutional uses of imperatives, but with descriptive uses. In these cases, the intuitive impact of "$A$ or $P$" is sufficiently captured in the discourse relation $require(a, p)$, and that is to deny a speech-act conjunction analysis and endorse a conditional analysis. For another, if we take parallel cases like (6) serious, the speech-act conjunction analysis is further discredited by the oddity of examples such as (7) where there is no sign of incongruity between disjuncts.

(7) a. ? It's a bad idea to invite Jason, or we'll have more beer for ourselves.
b. ? It's a bad idea to invite Jason, or he'll break Jane's heart (with the d.d.).

Examples (7) show that conditional uses of disjunction like (6), which are assumed to parallel IoDs, not only have a negative bias in terms of hearer-desirability in the second disjunct, but also a positive bias in terms of speaker-preferability in the first.

Therefore I suggest the following answer strategy for the IoD problem set for consideration: It is plausible and possible to assume that in "$A$ or $P$" the imperative form is only descriptively used and contributes to the expression of a natural conditional relation *require*$(a, p)$. Cases of pragmatic infelicity, like (7) or inappropriate instantiations of (1b), may then be seen as failures of the second disjunct to relate to the topic addressed in the first. Suppose that in "$A$ or $P$" the first disjunct introduces the action $a$ as a topic. Disjunction restricts the context of interpretation for the modal in the second disjunct to the non-topic worlds, yielding a conditional

relation $[\neg a]p$. We could then hypothesize that there is no discourse relation that relates $p$ via $[\neg a]p$ to the topic $a$ other than *require*$(a, p)$, which is to say that $p$ has to be undesirable. This is more than an ad hoc redescription of the original problem and it is also more than a bold non-existence claim, for it has indeed an intuitive basis. In particular, the answer to (Q3) from this perspective would be that there is no discourse relation *require'*$(a, p)$ defined as:

$$\textit{require'}(a, p) \quad \text{iff} \quad [\neg a]p \wedge \textit{cause}(\neg a, p) \ \wedge \ \neg \textit{undesirable}(p)$$

which would simply express a particularly awkward relation: 'omission of $a$ is necessary to bring about $p$'. This is an awkward relation in the light of the fact that it is supposed to be a relation about $a$ and not about $\neg a$. It is due to this relational lacuna that conditional uses of disjunction have the attested desirability biases and therefore instantiations of (1b) cannot request forbearance from the mentioned action.

## 5. Conclusion

This paper argued for a uniform and parsimonious conditional treatment of pseudo-imperatives. It was suggested that the persuading character of these sentences does not require the association of a particular illocutionary force with the featured imperative forms, but that these are merely descriptively used. An intuitive fact about relations in discourse was made responsible for the acceptability asymmetry in (1).

The most pressing open end is a compositional account of the assumed conditional readings. Remarks on where to look for such an account where already made for disjunction, but conditional readings of conjunctions are, as of yet, a particularly unpleasant gap in linguistic theory, which this paper did not attempt to bridge.[4]

## Bibliography

Clark, B.: 1993, Relevance and 'pseudo-imperatives', *Linguistics and Philosophy* 16, 79–121

Lascarides, A. and Asher, N.: 2004, Imperatives in dialogue, in P. Kuehnlein, H. Rieser, and H. Zeevat (eds.), *The Semantics and Pragmatics of Dialogue for the New Millenium*, Benjamins

Mann, W. and Thompson, S.: 1987, Rhetorical structure theory: A framework for the analysis of text, *IPRA Paper in Pragmatics* 1, 79–102

Mastop, R.: 2005, *What can you do? - Imperative Mood in Semantic Theory*, Ph.D. thesis, University of Amsterdam

Schwager, M.: 2004, *Don't be late or you'll miss the first slot*, unpublished manuscript, Talk held at NASSLLI, UCLA

van der Auwera, J.: 1986, Conditionals and speech acts, in E. C. Traugott, A. ter Meulen, J. S. Reilly, and C. A. Ferguson (eds.), *On Conditionals*, pp 197–214, Cambridge University Press

[4]Thank you: Paul, Robert, Samson, Peter & Monika and all who commented on my master thesis.

# AGENCY AND CASE: A LATTICE-BASED FRAMEWORK

SCOTT GRIMM

Institute for Logic, Language and Computation
University of Amsterdam
sgrimm@science.uva.nl

The typological literature has demonstrated that parameters such as agency, affectedness, and object individuation affect the realization of case-marking. The proposed analysis captures the specific contribution of such parameters, resulting in a model capable of explaining case alternations. A feature-based representation of agency properties is proposed, loosely based on Dowty's proto-role theory, but reformulated in terms of privative opposition and hierarchically organized via a lattice. Theoretical gains include wider empirical reach and greater simplicity, while practical results include a detailed analysis of the genitive/accusative alternation in Russian occurring with certain scope-ambiguous verbs, e.g. 'seek'.

Modulation of parameters such as agency, affectedness and object individuation are known to affect the realization of case-marking (Hopper and Thompson 1980). Yet, explicitly connecting individual parameters with the semantics of case alternations has largely proven elusive. Often, realizations of case cannot be attributed to one sole parameter, but arise only in the context of the interaction of several. These parameters are complex, and a large amount of typological work has sought to give them internal order on a universal basis, e.g., thematic, definiteness and animacy hierarchies. In the next section, I decompose the most fundamental parameters for argument structure, agency and affectedness, into feature-based representations, organized in section 2 into a lattice structure. This lattice in turn models argument structures. In section 3, the core semantics of a case is correlated with a region of the lattice, and by merging the lattice with the definiteness hierarchy, also decomposed in terms of features, I derive an account of the genitive/accusative alternation in Russian. In essence, the proposed framework yields both explanations of the semantic basis of case alternations and a more complete picture of how such parameters interact—in the meantime bringing the typological closer to the logical.

## 1. The Primitives of Argument Structure: Agency Properties

The parameters of agency and affectedness can be captured by a set of event-based properties entailed by the verb, inspired by the approach of Dowty 1991. However, the work of Dowty 1991 was elaborated taking the transitive situation as given,

therefore many of the proto-properties posited in Dowty 1991 are defined in terms of multiple participants, i.e., "causing an event or change of state in another participant". This assumption leads to difficulties in treating constructions beyond the typical transitive situation, such as the middle voice. Further, the properties of Dowty 1991 include the complex notions of 'affectedness' and 'causation' taken as primitive. Affectedness has long been noted not to be a binary concept, but a three-way distinction between unaffected, partially or totally affected. Causation is a complex notion, and in fact a composite one: implying at least two participants, and some sort of direct link between them. An increase in simplicity and empirical reach can be gained by reformulating the properties without reference to other participants and complex notions.

I use one set of properties, which can be conceived as of two types: one corresponding to the active ingredients of agentivity and the other to affectedness. The first set is comprised of the properties *volition*, *sentience*, *motion*, and *instigation*. The second set is cast in terms of *persistence*. Persistence is a two-tiered notion, for something can persist existentially, that is, its essence remains the same throughout the event/state, or it can persist qualitatively—i.e., it persists in all its particulars. Either of these can obtain at the beginning and/or the end of the event—in terms of features, we have the following set: *existential persistence (beginning)*, *existential persistence (end)*, *qualitative persistence (beginning)*, and *qualitative persistence (end)*. Establishing agency properties in this manner leads to two diametrically opposed classes in privative opposition, one a full agent possessing all the properties, and the other not entailing any, not even independent existence—e.g., arguments of negative existence statements or incorporated/cognate objects ("sing a song").

Affectedness can be reformulated as a lack of persistence during the event; further, this feature configuration is able to capture the different degrees of affectedness with respect to existence. Totally affected patients, e.g., verbs of destruction/consumption ('destroy', 'eat') entail that their object argument persists existentially at the beginning of the event, but not at the end. Patients which are partially affected (e.g., objects of verbs such as 'damage' or 'move') persist existentially throughout the event, but do not persist qualitatively, i.e., they are changed in some manner. Unaffected entities, most often agents, persist both existentially and qualitatively throughout the event. The opposition between agents and patients falls out from this feature system in that agents will possess total persistence along with a number of other agency properties while patients will generally possess no properties save initial persistence and possibly *existential persistence (end)*. The composite property of causation can be replaced by two more primitive ones: *instigation* and $-$ *persistence(end)* (either qualitative or existential). Causation, then, can be represented as a pair: (ArgX: $+$ *instigation*, ArgY: $-$ *persistence(end)*)[1].

[1] The results of this framework are conservative with respect to the gains of Dowty 1991. However, there is an empirical advantage to the proposed framework in that it can treat constructions outside the typical transitive situation—for instance, the middle voice in Ancient Greek, as shown in (1) (Lyons 1968):

## 2. Hierarchization of Agency Properties

The above has established a set of properties which make up a predicate's argument structure. Logical entailments among the eight features constrain the combinations possible. For instance, *volition* entails *sentience*, since only sentient beings are capable of volition, and *−existential persistence (end)* entails *−qualitative persistence (end)*, since if an entity does not exist at the end of the event, clearly none of its qualities do either. The remaining combinations can then be given greater structure. The sets of agency and persistence properties can be separately ordered by inclusion, giving rise to a lattice structure for each. The Cartesian product of the agency and persistence structures results in a larger lattice, shown in figure 1, referred to henceforth as the agency lattice. Note that the privative opposition is conspicuous in the structure: the highest node contains all the features (the full agent), the lowest contains none (event internal objects).

The parameter of object individuation—here, definiteness—is submitted to a similar treatment. Categories of definiteness can be reworked as a set of features (e.g., *referring* and *given*) ordered by inclusion, as in (1a), and corresponding to the definiteness hierarchy, seen in (1b).

(1) a. $\emptyset <$ referring $<$ referring, given
b. Non-Specific Indefinite$<$ Specific Indefinite$<$Definite

As opposed to the agency features which are predicate entailments, definiteness is endemic to the NP. Therefore, when a predicate's argument is instantiated with an NP, in this framework it is viewed as a merge of agency and individuation features.

## 3. The Genitive/Accusative Alternation in Russian

The agency features above are responsible for argument realization, i.e., which arguments are selected as subject, object, etc. One central function of case is to mark subjects (objects) as such, thus there is a necessary link between case and agency features. In languages with sufficiently expressive case systems, a subject (object) can be marked by a variety of cases. For instance, different sorts of object arguments

(1) Loúomai
Wash.1st.SING.MIDDLE
I am washing (myself).

Since there is only one participant in (1), the proto-properties "causing an event or change of state in another participant" and "causally affected by another participant" are undefined, and the only proto-properties that obtain are "undergoes a change of state", but this is not sufficient to differentiate subjects of verbs in the middle voice from subjects of verbs in the passive. What one would like to see is that the subject 'causally affects' himself, but this does not seem possible as long as causation is defined with respect to distinct participants. If instead, causation is defined, as discussed above, as a property for pairs (ArgX: + *instigation*, ArgY: − *persistence*) where ArgX and ArgY are not taken to be necessarily distinct, then this relation is applicable to the washer. For further details, see Grimm 2005.

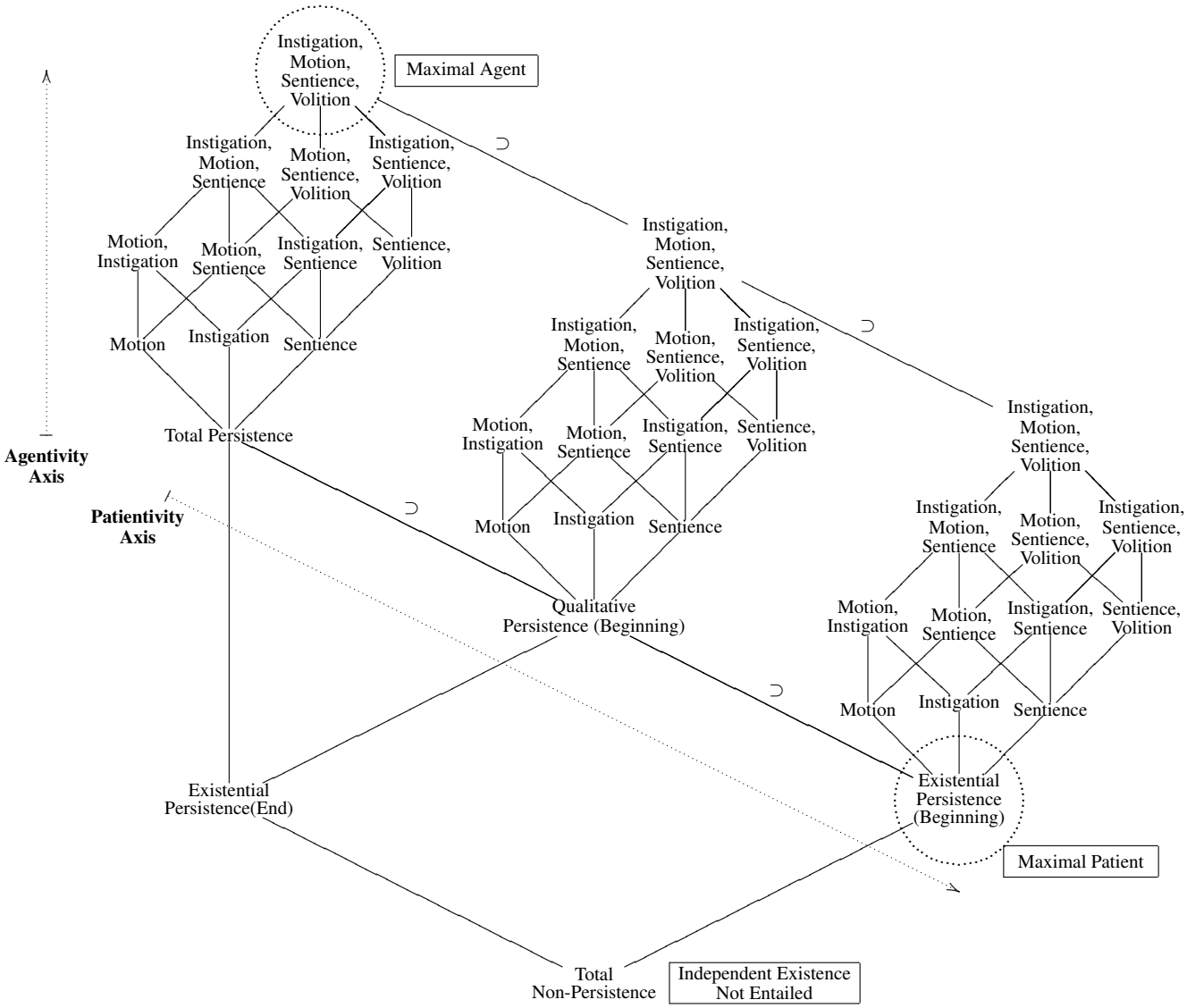



Figure 1: Transitivity Region

will be marked by different cases, e.g., recipients versus affected objects. Yet, these arguments can then be defined in terms of agency properties, and so one can define the core semantics of a case with respect to these agency features. A case-marker, then, can be seen as ranging over one or more (connected) node(s) in the lattice. Yet,

case assignment is not solely determined by argument structure, since the nominal instantiating the argument has its own semantic contribution which may or may not be consistent with the specifications of the argument structure. This is exemplified by the following alternation in Russian between genitive and accusative case on the object.

(2) RUSSIAN (Wierzbicka 1981)

a. Ivan ždet tramvaj-a
Ivan is-waiting-for tram-GEN
Ivan is waiting for a tram.

b. Ivan ždet tramvaj
Ivan is-waiting-for tram-ACC
Ivan is waiting for the/a certain tram.

Prima facie, (2) appears to mark (in)definiteness, and has been claimed as a form of differential object marking based on definiteness (Naess 2004). I claim that this is only true indirectly: it is a result of the interaction between the degree of definiteness of the NP and agency entailments on the verb. If definiteness were the only relevant factor, one would expect this alternation to apply generally; however, the above alternation is limited to verbs such as 'seek', 'await', 'want', 'fear'—verbs which are ambiguous between narrow- or wide-scope readings. While such verbs entail various agency properties in their subjects, they have no entailments for their objects, for to wait for a train does not necessitate that such a train exists. Note that with these verbs, the genitive marks the narrow-scope reading[2]. Historically, these verbs once marked their objects exclusively with the genitive case, and only gradually over the last century did this case alternation take hold.

Aside from this alternation, the more frequent use of the genitive as a verbal argument is to express lack of existence. For instance, "when an existential predicate is negated, the entity whose presence is denied is expressed in the genitive" (Timberlake 2004). Correlating this fact with the agency lattice, the governed genitive is used when existence of the object is not entailed, and therefore the genitive is associated with the lowest node of the lattice, *Total Non-Persistence*. In contrast, the accusative case marks objects of transitive clauses, and since these objects are generally affected in some way, they must be in existence before the onset of the event. Therefore, the region of the accusative covers at least the node *Existential Persistence (Beginning)*.

Definiteness enters the picture in the following way. Recalling that Ioup 1977 showed that referring arguments only have wide-scope readings, clearly an individuated (referring) object is not consistent with the semantics of the genitive, but instead forces a wide-scope reading, which *is* consistent with the semantics of the accusative.

[2]Russian morphology is well-known to be sensitive to wide- and narrow-scope interpretations. Dahl 1970 showed that the distribution of two suffixes of indefinites distinguished precisely these two readings.

The proposed framework captures this quite naturally. NPs which are specific or higher on the definiteness hierarchy, hence +*referring*, entail that the entity exists. Upon combining with the agency properties entailed by the predicate, the argument must then minimally possess the feature *existential (beginning)*, which locates the object in the region of the accusative case. If the NP is non-specific (non-referring), independent existence is not entailed, and it can remain on the lowest node of the lattice—but then this locates the object in the region of the genitive case. Therefore, definiteness is the crucial factor underlying this alternation, yet it is mediated by agency properties, which in turn explains the alternation's limited distribution.

## 4. Conclusion

A reformulation of the approach of Dowty 1991 with simpler primitives and in terms of privative opposition has led to greater empirical reach and a structured framework capable of accounting for the core semantics of case assignment. A principal advantage of this approach is its ability to unite multiple semantic parameters, as demonstrated by its account of the genitive/accusative alternation in Russian.

## Acknowledgements

Many thanks to my advisor Henk Zeevat and also to Andrej Malchukov for discussion.

## Bibliography

Dahl, O.: 1970, Some notes on indefinites, *Language* 46, 33–41

Dowty, D.: 1991, Thematic proto-roles and argument selection, *Language* 67, 547–619

Grimm, S.: 2005, The lattice of case and agentivity, *Master's thesis*, ILLC : University of Amsterdam

Hopper, P. J. and Thompson, S. A.: 1980, Transitivity in grammar and discourse, *Language* 56, 251–299

Ioup, G.: 1977, Specificity and the interpretation of quantifiers, *Linguistics and Philosophy* 1, 233–245

Lyons, J.: 1968, *Introduction to Theoretical Linguistics*, Cambridge

Naess, A.: 2004, *Transitivity*, Ph.D. thesis, Universiteit Nijmegen

Timberlake, A.: 2004, *A Reference Grammar of Russian*, Cambridge

Wierzbicka, A.: 1981, Case marking and human nature, *Australian Journal of Linguistics* 1, 43–80

# DYNAMIC WH-TERMS

ANDREAS HAIDA

Zentrum für Allgemeine Sprachwissenschaft (ZAS)
Berlin, Germany
haida@zas.gwz-berlin.de

The grammatical analysis of wh-questions presented in Groenendijk and Stokhof 1982 does not account for the close affinity between indefinite and interrogative pronouns. However, *wh*-terms can be treated in the same way as indefinites if existential quantification is dynamic. In this paper, the question denotations of G&S (1982) are reproduced in a dynamic framework in which *wh*-terms translate as existential GQs. In addition to this, the syntactic and semantic consequences for explaining the intervention effect in *wh*-questions are explored.

## 1. Introduction

Indefinite and interrogative pronouns are closely related in the majority of the world's languages.[1] This is exemplified in (1) and (2) with data from German and Lakhota, respectively.[2] As indictated, the *in-situ wh*-pronoun in both strings is ambiguous between an indefinite and an interrogative construal.[3]

(1) Wer hat was gekauft?
who has what bought
a. 'Who bought something?'
b. 'Who bought what?'

(2) šų́ka ki tákų yaxtáka he
dog the what bit Q
a. 'Did the dog bite something?'
b. 'What did the dog bite?'

If explanatory adequacy is to be achieved, a theory of interrogatives must therefore incorporate a compositional analysis of *wh*-questions in which *wh*-terms are treated essentially like indefinites. The question theory presented in Groenendijk and Stokhof 1982 is not equipped with such an analysis. Rather, this theory seems to entail that *wh*-terms are syncategorematic and that their closest categorematic counterparts are universally quantified terms.[4] However, with the development of dynamic semantics for natural language, an adequate conclusion could be reached: "Treating

[1] See Bhat 2000.
[2] For the latter, see Van Valin 1993.
[3] The two construals are disambiguated prosodically: if the *wh*-pronoun is accented, it must be construed as a question word, and otherwise as an indefinite. This seems to be a very general pattern.
[4] See Groenendijk and Stokhof 1982, p. 196 and 204f.

[*wh*-terms] like indefinites in a dynamic framework would mean translating them in terms of dynamic existential quantification. [...] [W]e might do so if for whatever reason this seems to be desirable after all." (G&S 1992, p. 122) In the following section, it is discussed how this proposal can be spelled out.

## 2. Question Denotations in a Dynamic Framework

According to G&S (1982), the denotation of an $n$-constituent interrogative is an index-dependent proposition that can be represented by a Ty2 expression of the form (3), where $\phi$ and $\psi$ are saturated relations $\beta(i)(x_1, \ldots, x_n)$ and $\beta(j)(x_1, \ldots, x_n)$, respectively.[5]

(3) $\lambda j(\lambda x_1 \ldots \lambda x_n.\phi = \lambda x_1 \ldots \lambda x_n.\psi)$

The gist of my account is that (4) denotes the same proposition as (3) if the existential quantifier and the biconditional connective are interpreted dynamically.

(4) $\lambda j(\exists x_1 \ldots \exists x_n.\neg\neg\phi \leftrightarrow \exists x_1 \ldots \exists x_n.\neg\neg\psi)$

For reasons of space, I will only give an intuitive argument for this equivalence. To simplify the discussion, assume that $\phi$ and $\psi$ do not have context change potential themselves. Then it must be shown that $\lambda x_1 \ldots \lambda x_n.\phi = \lambda x_1 \ldots \lambda x_n.\psi$ and $\exists x_1 \ldots \exists x_n.\phi \leftrightarrow \exists x_1 \ldots \exists x_n.\psi$ have the same truth conditions.

Consider first that an equation of the form $\lambda x_1 \ldots \lambda x_n.\phi = \lambda x_1 \ldots \lambda x_n.\psi$ can be equivalently[6] given as $\forall x_1 \ldots \forall x_n((\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi))$, and that universal quantification is distributive over conjunction:

$$\begin{gathered}\lambda x_1 \ldots \lambda x_n.\phi = \lambda x_1 \ldots \lambda x_n.\psi \\ \Leftrightarrow \\ \forall x_1 \ldots \forall x_n(\phi \rightarrow \psi) \wedge \forall x_1 \ldots \forall x_n(\psi \rightarrow \phi)\end{gathered}$$

In dynamic semantics, $\forall x(\Phi \rightarrow \Psi)$ is equivalent to $\exists x.\Phi \rightarrow \Psi$. Therefore, the following equivalence holds:

$$\begin{gathered}\forall x_1 \ldots \forall x_n(\phi \rightarrow \psi) \wedge \forall x_1 \ldots \forall x_n(\psi \rightarrow \phi) \\ \simeq \\ (\exists x_1 \ldots \exists x_n.\phi \rightarrow \psi) \wedge (\exists x_1 \ldots \exists x_n.\psi \rightarrow \phi)\end{gathered}$$

As can be easily verified, the last formula is true iff $\exists x_1 \ldots \exists x_n.\phi$ and $\exists x_1 \ldots \exists x_n.\psi$ have the same context change potential. It can then be asked for which connective '$\circ$' the formula $\Phi \circ \Psi$ is true iff $\Phi$ and $\Psi$ have the same context change potential. As is

[5] Cf. Groenendijk and Stokhof 1990, p. 1-9.

[6] In the following, the symbols '$\Leftrightarrow$' and '$\simeq$' are used to denote the equivalence of two formulas of static and dynamic logic, respectively.

argued below, the sought-after connective is the dynamic biconditional '$\leftrightarrow$'. That is, the following equivalence can be derived:[7]

$$(\exists x_1 \ldots \exists x_n.\phi \rightarrow \psi) \wedge (\exists x_1 \ldots \exists x_n.\psi \rightarrow \phi)$$
$$\simeq$$
$$\exists x_1 \ldots \exists x_n.\phi \leftrightarrow \exists x_1 \ldots \exists x_n.\psi$$

□

The dynamic biconditional is defined as given in (5) (where for each formula $\Phi$, $/\Phi/_{\mathcal{M},g}$ is the set of output assignments of $\Phi$ with respect to $\mathcal{M}$ and $g$).[8]

(5) **Definition:** $[\![\Phi \leftrightarrow \Psi]\!]_{\mathcal{M},g} = 1$ iff $/\Phi/_{\mathcal{M},g} = /\Psi/_{\mathcal{M},g}$

The reason for choosing this definition is that it implies that '$\leftrightarrow$' is the object-language counterpart of the metalanguage equivalence notion '$\simeq$' (as defined along the lines of G&S (1991)):

(6) **Fact:** $\Phi \simeq \Psi$ iff $\forall \mathcal{M} \forall g : [\![\Phi \leftrightarrow \Psi]\!]_{\mathcal{M},g} = 1$

### 2.1. The grammar of *wh*-questions

Due to the equivalence of (3) and (4), *wh*-terms can be treated in the same way as indefinites. That is, *wh*-terms can be translated as existential generalized quantifiers (see 7) if the semantic representation language is interpreted dynamically.[9]

(7) a. who$_k$, what$_k$ $\rightsquigarrow \lambda P.\exists x_k.P(i)(x_k)$
b. which$_k$ $\alpha \rightsquigarrow \lambda P.\exists x_k(\alpha'(i)(x_k) \wedge P(i)(x_k))$, where $\alpha \rightsquigarrow \alpha'$

Moreover, the interrogative complementizer Q can be given a unique translation:

(8) Q $\rightsquigarrow \lambda p \lambda j(p(i) \leftrightarrow p(j))$

On these assumptions, the denotation of a *wh*-question can be compositionally derived as exemplified in (9). On the basis of the syntactic structure (9a), the *de dicto* reading of *which girl which boy loves* is derived as indicated in (9b). Thereby, $\mathbf{Q} = \lambda p \lambda j(p(i) \leftrightarrow p(j))$ and $\mathbf{E}_k = \lambda P \lambda P'.\exists x_k(P(i)(x_k) \wedge P'(i)(x_k))$.

(9) (I wonder) which girl which boy loves
a. [ Q [ which$_2$ girl [ which$_1$ boy loves $t_{\text{which}_2\text{ girl}}$ ] ] ]
b. $\mathbf{Q}(\lambda i.\mathbf{E}_2(\text{girl}')(\lambda i \lambda v'.\mathbf{E}_1(\text{boy}')(\lambda i \lambda v.\text{love}'(i)(v, v'))))$

[7] There is reason to assume that the dynamic biconditional is externally dynamic. Therefore, the equivalence below should rather be written as truth-conditional equivalence (s-equivalence in the sense of Groenendijk and Stokhof 1991).
[8] Neither in Staudacher 1987 nor in Groenendijk and Stokhof 1991 is a dynamic biconditional defined.
[9] The expressions given below have the appearance of Ty2 expressions, but this is only for notational convenience. They are best to be considered as abbreviations for expressions that encode the notion of context change in the object language. See Muskens 1996.

Note in particular that both *wh*-phrases are interpreted in the position in which they occur in the overt syntactic structure: *which girl* is interpreted *ex situ* and *which boy in situ*.

## 3. The Intervention Effect

The dynamic treatment of *wh*-terms does not only account for the affinity between indefinite and interrogative pronouns but also explains another crosslinguistic phenomenon, namely the intervention effect in *wh*-questions. In Beck (to appear), the intervention effect is characterized as follows: Certain elements, so-called interveners, may not occur between a *wh*-phrase and its licensing complementizer (see 10). Interveners are focusing elements such as *only*, the sentence negation *not*, and quantifiers such as *most* and *never*.

(10) *[ Q [. . . [ intervener [. . . *wh*-phrase. . . ] ]. . . ] ]

The deviance of the constructions in (11) exemplifies this phenomenon.

(11) a. *mâymiikhray chôop ʔáan nagsii lêmnay (Thai)[10]
nobody like read book which
'Which books does nobody like to read?'
b. ??Wer hat niemandem was gezeigt? (German)[11]
who has nobody what showed
'Who showed what to nobody?'

In Honcoop 1996, it is observed that the expressions that induce the intervention effect[12] "all create so-called *inaccessible* domains for binding, i.e. an indefinite DP that occurs inside the syntactic scope of these expressions cannot bind a pronoun that occurs outside of their syntactic scope." (Honcoop 1996, p. 93) This is illustrated with the discourses in (12).[13]

(12) a. John did*n't* buy a car$_i$ (*n't* $>$ O). *It$_i$ was too expensive.
b. *Most students* bought a car$_i$ (S $>$ O). *It$_i$ was quite expensive.
c. John *never* bought a car (*never* $>$ O). *It$_i$ was too expensive.

The intervention effect therefore indicates that the relation between a *wh*-phrase and its licensing complementizer is anaphora like (in the sense that the context change brought about by the former is evaluated by the latter). Note that this is exactly

[9] = (1a) in Beck (to appear)
[10] = (21) in Beck (to appear)
[11] Cf. (11a) in Beck (1996).
[12] Honcoop 1996 is concerned with a variant of the intervention effect that does not fall under the description given above (at least superficially). However, Honcoop's analysis can be easily adapted to account for the intervention effect as it is conceived here.
[13] Cf. (13) and (16) in Honcoop 1996.

how this relation is analyzed here.[14] Therefore, the dynamic approach predicts that a *wh*-term cannot function as a question constituent in an intervention configuration. However, in its current form it fails to account for the deviance of the constructions in (11).

This becomes evident by considering an example: (11a) has the syntactic structure simplistically sketched in (13a) and its denotation is derived as specified in (13b). (Below, $\mathbf{E}_k = \lambda P.\exists x_k.P(i)(x_k)$ and $\mathbf{Q}$ is as defined before.) What can then be observed is that (13b) does not represent the extension of a *wh*-question, but of a yes/no question, namely the extension of (13c).

(13) a. [ Q [ nobody$_1$ [ read what$_2$ ] ] ]
b. $\mathbf{Q}(\lambda i.\neg\mathbf{E}_1(\lambda i\lambda v.\mathbf{E}_2(\lambda i\lambda v'.\text{read}'(i)(v,v'))))$
c. = Does nobody read anything? / Does somebody read something?

How can this result be interpreted? On the one hand, the derived effect (unavailability of a certain reading) clearly differs from the intervention effect (deviance). On the other, the distribution of the derived effect is the same as of the intervention effect. Therefore, the goal must be to strengthen the derived effect.

### 3.1. Non-interrogative indefinites

The above problem raises the deeper question of what distinguishes interrogative from non-interrogative indefinites. This question is answered as follows: Due to a morphosyntactic property, interrogative indefinites enter into a syntactic relation with an interrogative complementizer Q. As a consequence of this relation, interrogative indefinites share a syntactic index with Q. These indices are interpreted in such a way as to guarantee that the biconditional connective evaluates the context change of *all* and *only those* indefinites that are co-indexed with Q.

According to these assumptions, the questions *Who read what?* and *Who read something?* differ from each other in the way specified in (14) and (15). Thereby, $\mathbf{Q}^V = \lambda p\lambda j(p(i) \overset{V}{\leftrightarrow} p(j))$.

(14) a. [ Q$^{\{1,2\}}$ [ who$_1$ [ read what$_2$ ] ] ]
b. $\mathbf{Q}^{\{x_1,x_2\}}(\lambda i.\mathbf{E}_1(\lambda i\lambda v.\mathbf{E}_2(\lambda i\lambda v'.\text{read}'(i)(v,v'))))$

(15) a. [ Q$^{\{1\}}$ [ who$_1$ [ read something$_2$ ] ] ]
b. $\mathbf{Q}^{\{x_1\}}(\lambda i.\mathbf{E}_1(\lambda i\lambda v.\mathbf{E}_2(\lambda i\lambda v'.\text{read}'(i)(v,v'))))$

The relativized biconditional connective '$\overset{V}{\leftrightarrow}$' is defined in (16). By this definition, the context change brought about by non-coindexed (that is, non-interrogative) indefinites is ignored.

(16) $[\![\Phi \overset{V}{\leftrightarrow} \Psi]\!]_{\mathcal{M},g} = 1$ iff $/\Phi/^V_{\mathcal{M},g} = /\Psi/^V_{\mathcal{M},g}$, where $A^V = \{h_{|V} \mid h \in A\}$

[14] See Butler 2000 for a comparable approach which, however, remains largely inexplicit.

To rule out the intervention construction (17a), it must be assured that $\Phi \overset{V}{\leftrightarrow} \Psi$ is undefined if there is a variable in $V$ that is not subjected to an accessible modification in $\Phi$ and $\Psi$.

(17) a. [ $Q^{\{2\}}$ [ nobody$_1$ [ read what$_2$ ] ] ]
b. $\mathbf{Q}^{\{x_2\}}(\lambda i.\neg\mathbf{E}_1(\lambda i\lambda v.\mathbf{E}_2(\lambda i\lambda v'.\text{read}'(i)(v,v'))))$

This is achieved by partializing the evaluation contexts (left out for reasons of space).

**Acknowledgements**

Thanks to Sigrid Beck, Regine Eckardt, Cornelia Endriss, Gisbert Fanselow, Stefan Hinterwimmer, Manfred Krifka, Chris Potts, Uli Sauerland, and Peter Staudacher for helpful discussions.

**Bibliography**

Beck, S.: to appear, Intervention effects follow from focus interpretation, *Natural Language Semantics*

Bhat, D. N. S.: 2000, The indefinite-interrogative puzzle, *Linguistic Typology* 4, 365–400

Butler, A.: 2000, Semantic control and intervention, in C. Pilière (ed.), *Proceedings of the ESSLLI-2000 Student Session*, pp 45–54

Groenendijk, J. and Stokhof, M.: 1982, Semantic analysis of *wh*-complements, *Linguistics and Philosophy* 5, 175–233

Groenendijk, J. and Stokhof, M.: 1990, *Partitioning Logical Space*, Annotated handout ESSLLI 1990, Leuven

Groenendijk, J. and Stokhof, M.: 1991, Dynamic predicate logic, *Linguistics and Philosophy* 14, 39–100

Groenendijk, J. and Stokhof, M.: 1992, A note on interrogatives and adverbs of quantification, in C. Barker and D. Dowty (eds.), *Proceedings from SALT II*, pp 99–124, The Ohio State University, Columbus, Ohio

Honcoop, M.: 1996, Towards a dynamic semantics account of weak islands, in T. Galloway and J. Spence (eds.), *Proceedings from SALT VI*, pp 93–110, Cornell University, Ithaca

Muskens, R.: 1996, Combining Montague semantics and discourse representation, *Linguistics and Philosophy* 19, 143–186

Staudacher, P.: 1987, Zur Semantik indefiniter Nominalphrasen, in B. Asbach-Schnitker and J. Roggenhofer (eds.), *Neuere Forschungen zur Wortbildung und Historiographie der Linguistik*, pp 239–258, Narr, Tübingen

Van Valin, Robert D., J.: 1993, A synopsis of role and reference grammar, in R. Van Valin (ed.), *Advances in Role and Reference Grammar*, pp 1–164, Benjamins, Amsterdam

# CONTRASTIVES AND GRICEAN PRINCIPLE

YURIE HARA

Department of Linguistics
University of Delaware/University of Massachusetts, Amherst
yhara@udel.edu

It has been observed that contrastive-markings in various languages are associated with uncertainty implicatures. However, a sentence can be contrastive-marked even when the speaker has a complete answer to the question, as long as one of the alternatives have an opposite value from the rest. Following the analyses by Spector 2003 and Schulz and van Rooij (in press) on exhaustivity and the Gricean Principle, this paper claims that Contrastive-marking presupposes that the speaker's knowledge is not maximal.

## 1. Introduction

Contrastive meaning can be represented by prosody as in German (Topic-Focus contour) and English (B-accent) and also by prosody and morphology as in Japanese (*-wa*) and Korean (*-nun*). Previous analyses claim that their contrastive meanings come from an *uncertainty* implicature as witnessed in (1-a) and (2-b). However, uncertainty alone does not correctly characterize all the properties of Contrastives. This paper elaborates an analysis that connects the phenomena to a more general pragmatic principle, i.e. the Griean principle, rather than stipulating that Contrastives induce a certain implicature.

(1) /ALLE Politiker sind NICHT\ korrupt
all politicians are not currupt
a. 'It is not the case that all politicians are corrupt.'($\neg\forall$)
(Possibly, some are corrupt.)
b. *'No politicians are corrupt.' (*$\forall\neg$)
(No implicature: unavailable reading for (1)) (German; Büring 1997)

(2) a. Who passed the exam?
b. MARY-wa ukat-ta
Mary-Con pass-Past
'[Mary]$_{Con}$ passed.'
c. MARY-ga ukat-ta.
Mary-Nom pass-Past
'Mary passed.' (exhaustive answer) (Japanese)

## 2. Büring 1997 and Hara 2004

Büring 1997 claims that a contrastive-marked sentence implicates there exist some unanswered questions: "If a sentence S with a Topic accent is uttered given some Context CX, and there is no disputable Residual Topic the sentence establishes, the utterance of S in CX is infelicitous." Similarly to Büring, Hara 2004 claims that a contrastive-marked sentence presupposes that there exist some stronger scalar alternative to the assertion, and it implicates that it is possible that the stronger alternative is false.

(3) CONTRASTIVE($w$)($x$)(B)(F) ($w$: world variable, $x$: speaker or attitude-bearer, F: contrastive-marked element, B: background, C: common ground)
   a. asserts: B(F)($w$)
   b. presupposes: $\exists$F′[[F′ $\in$ $ALT_C$(F)] & [B(F′) $\Rightarrow$ B(F)]& [B(F) $\not\Rightarrow$B(F′)]]
   c. implicates: $\exists w' \in \min_w[w' \in \text{Dox}_x(w)]$ : $\forall$F′[[F′ $\in$ $ALT_C$(F)] & [B(F′) $\Rightarrow$ B(F)]& [B(F) $\not\Rightarrow$B(F′)]][B(F′)($w'$)=0]

## 3. Contrastives can be used with a complete answer

However, questions can be completely resolved as in (4-a) and (4-c). What is prohibited is to have positive answers for **all** the alternatives as in (4-b) and (4-d).

(4) Among John, Mary and Bill, who came to the party?
   a. /JOHN und MARIA\ sind gegangen, (aber) /BILL ist NICHT\ gegangen.
      John and Maria are gone, but Bill is not gone
   b. */JOHN und MARIA\ sind gegangen, (aber) /BILL ist GEGANGEN\.
   c. John-to Mary-**wa** ki-te/takedo, Bill-**wa** ko-nakat-ta.
      John-and Mary-Contrastive come-and/Past.but, Bill-Contrastive come-Neg-Past
   d. *John-to Mary-**wa** ki-te/takedo, Bill-**wa** ki-ta.

## 4. Exhaustivity (Spector, 2003; Schulz and van Rooij, (in press))

Notice that the induced implicatures are very similar to conversational scalar implicatures of Grice. In recent literature of Gricean framework, scalar implicatures are derived in two steps. Especially, Spector 2003 and Schulz and van Rooij (in press) derive scalar implicatures from exhaustivity. For example, as the first step, Schulz and van Rooij define the order of knowledge as follows: "a speaker has more knowledge about $P$ if she knows of more individuals that they have property $P$." In other words, the crucial assumption here is that in the case where the speaker

knows of some individuals **not** having property $P$, it is not counted as the speaker's knowledge with respect to $P$, as they note in their footnote 45 [p. 41]. Given this assumption, Schulz and van Rooij (in press) formulates the Gricean Principle:[1]

(5) Interpreting according to the Gricean Principle (Schulz and van Rooij (in press))
Let $A$ be an answer given to a question with question-predicate $P$ in context $C =< W, R >$. We define the pragmatic interpretation $grice^C(A, P)$ of $A$ with respect to $P$ and $C$ as follows:
$grice^C(A, P) =_{def} \{w \in [\mathbf{K}A]^C | \forall w' \in [\mathbf{K}A]^C : w \preceq_{P,A} w'\}$

Hence, the Gricean Principle gives a primary (weak) implicature, the speaker does not have knowledge whether the property in question holds for the rest of the alternative individuals.

Second, the assumption that the speaker is competent gives a secondary (strong) implicature, the speaker knows that the property does not hold for the alternative individuals.[2] Hence, we obtain a welcome exhaustive interpretation by maximizing the speaker's competence.

(6) Adding Competence to the Gricean Principle (Schulz and van Rooij (in press))
$eps^C(A, P) =_{def} \{w \in grice^C(A, P) | \forall w' \in grice^C(A, P) : w \not\sqsubset_{P,A} w'\}$
$= \{w \in [\mathbf{K}A]^C | \forall w' \in [\mathbf{K}A]^C :$
$w \preceq_{P,A} w' \wedge (w \cong_{P,A} w' \rightarrow w \not\sqsubset_{P,A} w')\}$

## 5. Contrastives Necessarily Limit Competence

As mentioned earlier, primary weak implicatures are very similar to the implicatures associated with Contrastive-marking. In (2), (2-b) indicates the speaker is not sure about others.

Hence, I propose that the function of Contrastive is to indicate that the speaker has a limited competence with respect to the question predicate.[3] Now, I posit the following interpretation of Contrastive. Here, I go back to the notations in structure meaning approach I used in Hara 2004. The background B and the assertion B(F)

[1] $w_1 \preceq_{P,A} w_2$ means that the speaker in $w_2$ is more or equally knowledgeable compared to the speaker in $w_1$. See Schulz and van Rooij (in press) for a more precise definition.

[2] $w_1 \sqsubseteq_{P.A} w_2$ roughly means that the speaker in $w_2$ is more or equally competent than the speaker in $w_1$. $w \cong_{P,A} w'$ roughly that means the knowledge of the speaker in $w$ is equal to that in $w'$. See Schulz and van Rooij (in press) for more precise definitions.

[3] Schulz and van Rooij (in press) also mentions this intuition at the end of their section 7 [p. 49]: "the answerer can cancel this additional assumption by either mentioning that she is not competent or simply deviating from the standard form of answering a question (by using negation, special intonation, etc.). In this way we can correctly predict the weakening of exhaustive interpretation to 'limited-competence' inferences for such answers."

correspond to the question predicate $P$ and the answer $A$ respectively in Schulz and van Rooij (in press):

(7) Interpreting a sentence with Contrastive-marking
CONTRASTIVE(B(F))
presupposes: $eps^C$(B(F), B) $\neq$ $grice^C$(B(F), B)
implicates: $grice^C$(B(F), B) $= \{w \in [\mathbf{K}(\text{B(F)})]^C | \forall w' \in [\mathbf{K}(\text{B(F)})]^C : w \preceq_{\text{B,B(F)}} w'\}$

Contrastive lexically specifies that the speaker's competence is minimal and prevents the interpreter to draw the exhaustive interpretation (i.e. the secondary strong implicature). To illustrate, let us take the example (2) assuming that there are only two individuals, Peter and Mary. (9) is the interpretation of (2-b).

(8) a. B= $\lambda x.\mathbf{passed}(x)$
b. F= $m$

(9) CONTRASTIVE($\mathbf{passed}(m)$):
implicates:$grice^C(\mathbf{passed}(m), [\lambda x.\mathbf{passed}(x)])$
$= \{w \in [\mathbf{K}(\mathbf{passed}(m))]^C | \forall w' \in [\mathbf{K}(\mathbf{passed}(m))]^C : w \preceq_{[\lambda x.\mathbf{passed}(x)],\mathbf{passed}(m)} w'\}$
$\models \neg\mathbf{K}(\mathbf{passed}(p))$

Just like the primary implicature computation, the result of (9) entails that the speaker considers both the possibility where Peter passed and the possibility where Peter didn't pass. Hence, it entails that the speaker doesn't know that Peter passed, which seems to be a desired interpretation for (2-b).

Moreover, Contrastive-marking not only generates implicatures when possible, but **always** generate implicatures. Therefore, Contrastive-marking is possible only in the environment that the speaker's knowledge is limited. Namely, there must be an effect by limiting the competence (the presuppositional requirement in (7)). This explains the infelicity of (10).

(10) #ZEN'IN-wa kita.
Everyone-Con came
'[Everyone]$_{Con}$ came.'

Knowing that 'Everyone came.' is true entails knowing that all the individuals are in the extension of the property $\lambda x \in D_e.x$ came. Applying $eps^C$ does not affect the interpretation since the assertion itself implies that the speaker has a maximal knowledge with respect to the property; hence the speaker is maximally competent, which is not compatible with the presupposition of Contrastive-marking.

Now, how does (7) overcome the problem presented in section 3.? Let us illustrate with the following simplified examples assuming again that we are only considering two individuals, Peter and Mary.

(11) Who passed the exam?

a. MARY-wa ukat-te/takedo, PETER-wa ukara-nakat-ta
Mary-Con pass-and/Past.but, Peter-Con pass-Neg-Past
'[Mary]$_{Con}$ passed and/but [Peter]$_{Con}$ didn't pass.'

b. *MARY-wa ukat-te/takedo, PETER-wa ukat-ta
Mary-Con pass-and/Past.but, Peter-Con pass–Past
'[Mary]$_{Con}$ passed and/but [Peter]$_{Con}$ passed.'

As we have seen earlier, the first conjunct of (11-a) '[Mary]$_{Con}$ passed.' implicates that $\neg$**K**$B_1(p)$ 'the speaker doesn't know that Peter passed.' This is compatible with the assertion of the second conjunct. Remember that knowing that an individual does not have a property B does not count as knowledge.

(12) a. $B_1$= $\lambda x$.**passed**$(x)$
b. $F_1$= $m$

(13) a. Assertion of the first conjunct of (11-a) entails:
**K**($B_1$($F_1$)) (=**K**(**passed**$(m)$))
b. Interpretation of CONTRASTIVE($B_1$($F_1$)): $\neg$ **K** $B_1$($p$)
c. Assertion of the second conjunct of (11-a) (in terms of $B_1$) entails:
**K**($\neg B_1(p)$) (**K**(=$\neg$**passed**$(p)$))
d. $\neg$ **K** $B_1$($p$) and **K**($\neg B_1(p)$) are compatible.

Similarly, the second conjunct '[Peter]$_{Con}$ didn't pass.' has the following values for $B_2$ and $F_2$.

(14) a. $B_2$= $\lambda x.\neg$**passed**$(x)$
b. $F_2$= $p$

Hence, CONTRASTIVE($\neg$**passed**$(p)$) brings an interpretation 'the speaker doesn't know that Mary didn't pass.' Again, this is compatible with the assertion of the first conjunct, which can be translated into 'Mary does not have the property of $B_2$ (non-passing).'

(15) a. Assertion of the second conjunct of (11-a) entails:
**K**($B_2$($F_2$)) (=**K**($\neg$**passed**$(p)$))
b. Interpretation of CONTRASTIVE($B_2$($F_2$)): $\neg$ **K** $B_2$($m$)
c. Assertion of the first conjunct of (11-a) (in terms of $B_2$) entails:
**K**($\neg B_2(m)$) (=**K**($\neg\neg$**passed**$(m)$))
d. $\neg$ **K** $B_2$($m$) and **K**($\neg B_2(m)$) are compatible.

On the other hand, (11-b) raises incompatibility between the contrastive interpretation of the first conjunct and the assertion of the second conjunct. Unlike the case of (11-a), the second conjunct of (11-b) tells that the speaker knows that $p$ has the property $B_1$, which contradicts the interpretation of the CONTRASTIVE($B_1$($F_1$)). The

same result comes out for the second conjunct (omitted for space reasons).

(16) a. Assertion of the first conjunct of (11-b) entails:
**K**($B_1(F_1)$) (=**K**(**passed**$(m)$))
b. Interpretation of CONTRASTIVE($B_1(F_1)$): $\neg$ **K** $B_1(p)$
c. Assertion of the second conjunct of (11-b) (in terms of $B_1$) entails:
**K**($B_1(p)$) (=**K**(**passed**$(p)$))
d. $\neg$ **K** $B_1(p)$ and **K**($B_1(p)$) are incompatible!

In summary, in order to account for the contrast between (11-a) and (11-b), it is necessary to assume the order of knowledge proposed by Schulz and van Rooij (in press), i.e. to exclude from the speaker's knowledge the case where the speaker knows of an individual **not** having the property.

## 6. Conclusion

Contrastive-marking a sentence indicates that the speaker has a limited competence with respect to the property in question; and therefore, it lexically induces Gricean primary implicatures. Also, the order of knowledge proposed by Schulz and van Rooij (in press) is crucial to make the correct predictions. Especially, the speaker can use Contrastive-marking even when the speaker has answers for all the individuals as long as there is a contrast among them.

## Acknowledgements

I am grateful for comments and discussions from Jan Anderssen, David Beaver, Benjamin Bruening, Daniel Büring, Gennaro Chierchia, Kai von Fintel, Danny Fox, Pauline Jacobson, Angelika Kratzer, David Oshima, Anna Papafragou, Paul Portner, Chris Potts, Maribel Romero, Florian Schwarz, Satoshi Tomioka, Katsuhiko Yabushita and anonymous reviewers. All errors are my own responsibility.

## Bibliography

Büring, D.: 1997, The great scope inversion conspiracy, *Linguistics and Philosophy* 20, 175–194

Hara, Y.: 2004, Scope inversion in Japanese: Contrastive topic require implicatures, in *Japanese/Korean Linguistics 13*, CSLI Publication

Schulz, K. and van Rooij, R.: in press, Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation, *Linguistics and Philosophy*

Spector, B.: 2003, Scalar implicatures: exhaustivity and gricean reasoning, in B. ten Cate (ed.), *Proceedings of the ESSLLI'03 student session*

# INFERENCE, ELLIPSIS AND DEACCENTING

DANIEL HARDT

Computational Linguistics
Copenhagen Business School
dh@id.cbs.dk

It is proposed that inference is available for ellipsis interpretation, only when triggered by a semantically visible violation. This captures well-known observations about available inferences in VP ellipsis, and new observations about unavailable inferences in sluicing. Furthermore, it is argued that inference in deaccenting is governed by the same triggering constraint.

## 1. Introduction

Paul Grice, in a famous pronouncement, enjoins language users to "avoid unnecessary prolixity" (Grice 1975). Necessary or not, redundant material is ubiquitous in natural language, its presence typically signalled by deaccenting or ellipsis. Naturally enough, deaccenting and ellipsis are infelicitous if the relevant material is not understood to be redundant. This is the minimal condition on deaccenting and ellipsis, and, ideally, this requirement would suffice to uniformly characterize its distribution.

One apparent problem with this ideal view involves inference: while inference is clearly involved in determining redundancy with *deaccenting*, this does not appear to be the case with *ellipsis*, as illustrated by the following examples from Rooth 1992 (*italics* indicates deaccenting.):

(1) First someone told Mary about the budget cuts, then SUE *heard about them*.

(2) First someone told Mary about the budget cuts, then SUE did.

The Verb Phrase *heard about them* is deaccented felicitously in (1), even though it has not appeared previously. Rooth argues that this relies on the inference *if someone tells Mary about budget cuts then Mary hears about them*. This inference is not available for (2), which does not have the reading *Sue heard about them*. Based on such facts, Rooth argues that deaccenting and ellipsis cannot be unified.

In fact, there is an early argument, due to Webber 1978, that inference is indeed required for ellipsis interpretation. Rooth does not address this argument, which

has been ignored in much of the subsequent theoretical literature. Certain recent authors have echoed Webber's argument (Fox 2000; Merchant 2001). Like Webber, these authors point to examples of ellipsis where the correct interpretation requires an inference.

These authors argue convincingly that there is good reason to take inference seriously in ellipsis. I will argue, however, that what is most intruiging about inference is its *absence* – it has not been previously observed that there are various cases of ellipsis and deaccenting where inference is systematically unavailable. I will argue that this reflects a fundamental, general condition on inference in interpretation: it must be *triggered* by a violation. I will argue, furthermore, that inference in deaccenting is subject to the same constraint. The evidence, then, points to a unified theory of deaccenting and ellipsis; appearances to the contrary can be traced to the simple fact that deaccenting can more easily trigger inference, because it involves overt lexical material.

In what follows, I first present the argument for inference in ellipsis, as given by Webber. Next, I present various cases where inference is unavailable in ellipsis: these primarily involve case-matching effects in sluicing. I then show that inference is also unavailable for deaccenting, in an example involving scope parallelism. I give a general statement of the triggering condition, and I argue that this uniformly accounts for inference in ellipsis and deaccenting.

## 2. Inference in Ellipsis: the Argument

Webber 1978 argues that inference is sometimes required in ellipsis interpretation, as in the following example of Verb Phrase (VP) Ellipsis:

(3) Irv and Mary wanted to dance together, but Mary couldn't, because her husband was there.

Webber notes that the desired reading is not *Mary couldn't dance together*, but rather, *Mary couldn't dance with Irv*. She argues that this results from the following inference:

(4) Irv and Mary wanted to dance together $\Rightarrow$ Mary wanted to dance with Irv.

One might wonder if inference is indeed required: an alternative explanation is that what is elided is not the entire VP *dance together*, but just the verb *dance*, giving *Irv and Mary wanted to dance together, but Mary couldn't dance*. This is then interpreted as *dance with Irv* because of context.

While this appears to be plausible alternative account for (4), this will not account for examples like (5), as pointed out in Hardt 1993:

(5) Martha and Irv had planned to nominate each other, but Martha couldn't, because of her political obligations.

Here, the reading is "Martha couldn't nominate Irv". Unlike *dance*, the verb *nominate* cannot appear without its object.

Webber notes that "there seem to be no hard and fast rules delimiting the class of productive inferences relative to verb phrase ellipsis." (p. 4-38) However, there is a crucial factor in example (4) which is not noted by Webber – the interpretation without inference is unacceptable, involving, as it does, an agreement violation between the singular subject *Mary* and the VP *dance together*.

This fact is the key to solving the problem posed by Webber, about how to delimit the class of available inferences in ellipsis: only inferences triggered by violations are possible. To show this, we examine cases where inferences are systematically unavailable.

## 3. Missing Inferences

We have seen that inference is possible in ellipsis when triggered by a violation. In this section, I will argue that inference is not available in the absence of a violation. Consider the following variant of (4), where the VP ellipsis has a plural subject:

(6) Irv and Martha wanted to dance together, but Tom and Susan didn't want to.

Here, the default, or non-inference reading is acceptable: *Tom and Susan didn't want to dance together*. In this case, the inference is no longer available: this variant cannot mean *Tom and Susan didn't want to dance with Irv*.

We turn now to sluicing. As first observed by Ross 1967 sluicing is subject to a case-matching requirement, as illustrated by the following example:

(7) Er will jemandem schmeicheln, aber sie wissen nicht
He wants someone.DAT flatter, but they know not
wem.
who.DAT.
He wants to flatter someone, but they don't know whom

(8) Er will jemandem loben, aber sie wissen nicht wen.
He wants someone.ACC flatter, but they know not who.ACC.
He wants to flatter someone, but they don't know whom

As Ross points out, the wh-word in (7) must take dative case, which is the case assigned by *flatter*. In (8), the wh-word is required to take accusative case, which is the case assigned by *praise*.

This *case matching* constraint in sluicing is stated by Merchant 2001 as follows:

- *The sluiced wh-phrase must bear the case that its correlate bears.*

What has not previously been noticed is that case matching fails to interact with inference. This is illustrated by (9):

(9) John traf jemanden, aber ich weiss nicht wen [John traf x] /*wer [traf John].
John met someone, but I don't know who [John met x] / who [met John].

Here, the wh-word must have accusative case ("wen"), the nominative "wer" gives rise to ill-formedness. What is interesting here is that "met" is a symmetric predicate: "John met someone" is semantically equivalent to "Someone met John", so on Merchant's semantic condition, the nominative "wer [traf John]" should also be permitted.

I conclude that inference is possible in ellipsis if and only if there is a semantically visible violation. In the following section, I give evidence that the same is true with deaccenting.

## 4. Missing Inferences in Deaccenting

Consider the following sentences:

(10) A doctor saw every patient. A NURSE *saw every patient, too*.

(11) A doctor saw every patient. A NURSE *saw many patients, too.*

(*Italics* indicates deaccenting.) It is widely acknowleged that examples like (10) are subject to a scope parallelism constraint – that is, *a doctor* takes wide scope in the first sentence only if *a nurse* also takes wide scope in the second. A similar scope parallelism effect can be observed in (11). Note that in this case, the deaccented material does not match – *every* has changed to *many*. This is the result of the following intervening inference, which I call the *every-many* inference:

(A doctor x) (every patient y) x saw y. → (A doctor x) (many patient y) x saw y.
(This follows if one assumes that there are many patients.)

The situation here is exactly analogous to the problem with case-matching: the possibility of inference threatens to undermine the scope parallelism constraint, since there is an inferential relation between the two scopes of *exists* and *every*. Consider the following LF representations for (10):

(12) a. (A doctor x) (every patient y) x saw y.
b. (every patient y) (A NURSE x) x saw y.

The intention of the scope parallelism constraint is to rule out this representation. But parallelism could be satisfied by the following intervening inference: (A doctor x) (every patient y) x saw y → (every patient y) (A doctor x) x saw y. To maintain

scope parallelism, such scope-shifting inferences must be ruled out, although in other cases, like (11), scope parallelism determination must permit intervening inferences like the *every - many* inference. How is the scope-shifting inference ruled out, while the *every-many* inference is permitted?

In previous work Hardt 2005 I have suggested that inferences be restricted by a *submodel* constraint. However, it is not clear that there is evidence for such a constraint. What is crucial in the current example is that the scope shifting inference is not required to satisfy parallelism. Thus the triggering constraint rules it out.

I turn now to a general statement of the triggering constraint.

## 5. Triggering Constraint

I propose the following general account of inference and interpretation: for a given discourse D, we produce a *default* LF $L$. If L violates no *semantically visible* constraints, it is the preferred interpretation. If L does violate one or more constraints, inferences can be performed to derive an alternative interpretation $L'$. $L'$ is a potential interpretation of D if it avoids the constraint violations. If there are several such alternatives, those LF's *closest* to $L$ are preferred. The closeness relation is defined with respect to entailment; if A entails B and B entails C, then B is closer to A than C.

I will not give any general characterization of *semantically visible* constraints: in this paper, I will rely on what I take to be completely standard conceptions, namely, that the parallelism requirement is semantically visible, as are agreement violations such as observed in (3). On the other hand, the *case-matching* constraints in sluicing are not semantically visible.

The triggering constraint proposed here has roots in the proposal of Fox 2000, where it is proposed that inference in ellipsis interpretation must be triggered. There are important differences, however. Most important is the triggering condition – on my proposal, semantically visible violations trigger inference, while for Fox the triggering condition involves a mismatch of lexical items. That is, Fox claims that inference is triggered by deaccented overt material in the ellipsis clause that is not present in the antecedent clause. This proposal is far less general than the current proposal, which is simply that semantically visible violations trigger inference. In my view, Fox's proposal lacks the intuitive motivation of my proposal. Furthermore, many cases of triggered inferences, including all those discussed in this paper, are beyond the scope of Fox's proposal.

## 6. Conclusions

Ellipsis and deaccenting signal redundancy, and thus they are subject to a requirement that they apply to material that is understood as redundant. On the one hand, it is natural to suppose that the determination of redundancy proceed with the help

of inferential reasoning. On the other hand, inferences can be computationally expensive operations, and dependence on inference might be difficult to reconcile with the demands of online processing. The evidence presented in this paper leads to the conclusion that inference is generally available, but in a very limited way – it is only available when triggered by a semantically visible violation.

## Acknowledgements

This work is supported by a grant from the Danish Research Council for the Humanities.

## Bibliography

Chung, S.: 2005, *A Look at Sluicing in Chamorro (and Germanic)*, Handout, University of Santa Cruz

Fox, D.: 2000, *Economy and Semantic Interpretation*, MIT Press

Grice, P.: 1975, Logic and conversation, in P. Cole and J. Morgan (eds.), *Syntax and Semantics 3*, pp 41–58, Academic Press

Hardt, D.: 1993, *Verb Phrase Ellipsis: Form, Meaning, and Processing*, Ph.D. thesis, University of Pennsylvania

Hardt, D.: 2005, Natural language inference as triggered submodel search, in *Proceedings, FLAIRS05*, Clearwater Beach, Florida

Merchant, J.: 2001, *The syntax of silence: Sluicing, islands, and identity in ellipsis*, Oxford

Rooth, M.: 1992, Ellipsis redundancy and reduction redundancy, in *Proceedings of the Stuttgart Ellipsis Workshop*, Stuttgart, Germany

Ross, H.: 1967, *Constraints on Variables in Syntax*, MIT Department of Linguistics and Philosophy

Webber, B. L.: 1978, *A Formal Approach to Discourse Anaphora*, Ph.D. thesis, Harvard University, (Published 1979 by Garland Publishing, New York)

# ASYMMETRIES IN LANGUAGE USE REVEAL ASYMMETRIES IN THE GRAMMAR

PETRA HENDRIKS[1], HELEN DE HOOP[2] & MONIQUE LAMERS[2]

[1]CLCG
University of Groningen
P.Hendriks@rug.nl

[2]CLS
Radboud University Nijmegen
{H.deHoop, M.Lamers}@let.ru.nl

According to the classical view, the nature of the grammar is independent of its use. A puzzle for this view is the observation that in language acquisition production sometimes precedes comprehension. Children who use subject-object status to determine word order in production not necessarily use word order to determine subject-object status in comprehension. On the basis of results from first language acquisition as well as adult sentence processing, we show that the role of grammar in production can be different from the role of the same grammar in comprehension.

## 1. Asymmetry in Language Acquisition

If children produce a particular form correctly, they seem to know the relevant grammatical rules. But if they know the relevant grammatical rules, they should be able to use these rules in comprehension too. However, children's performance in production and comprehension sometimes diverges, suggesting an asymmetric grammar. The aim of this paper is to provide additional evidence for such an asymmetric grammar and discuss the properties of this grammar.

A remarkable asymmetry in language acquisition has been found with word order in English. In an experiment with 15 children (1;8-2;8), Chapman and Miller (1975) found that in production these children tend to preserve subject-object order. In 83.7% of the cases with two animate actors, children produced forms such as "boy hit girl" or "hit girl", rather than "girl hit boy" or "hit boy", when having watched the action of a boy doll hitting a girl doll. However, these same children, when tested on the same type of sentences in a comprehension experiment, significantly less often (in 66.5% of the cases) used word order information as a cue to subject-object status. These children frequently demonstrated the action expressed by the sentence *The boy is hitting the girl* with the girl doll hitting the boy doll. Apparently, for a young child this sentence can mean both "boy hit girl" and "girl hit boy". A similar asymmetry between production and comprehension was found by McClellan, Yewchuk and Holdgrafer (1986) in their replication of

Chapman and Miller's experiment. These observations are further supported by the frequent occurrence of inversion errors in acquisition experiments on anaphora comprehension in e.g. Dutch, reported by Koster (1993).

To account for the observed production/comprehension asymmetry within the grammar, a theory of grammar is required that distinguishes between the hearer's task and the speaker's task. Optimality Theory (henceforth OT, Prince and Smolensky 2004) provides such a theory. In OT, it is assumed that language users select the best output for a given input on the basis of optimization over an ordered set of violable constraints. In production, the input is a meaning and the output the optimal form for that meaning. In comprehension, the input is a form and the output the optimal meaning for that form. In section 2 we show how OT can account for the pattern found by Chapman and Miller.

## 2. PRECEDENCE and PROMINENCE

Several off-line and on-line experiments in English and Dutch have proven that besides word order, animacy is an important source of information in comprehension (e.g., McDonald 1987). In their study of animacy in sentence comprehension, de Hoop and Lamers (to appear) propose a set of five violable constraints that serve the purpose of distinguishing the subject from the object in transitive sentences. The following two constraints seem to be relevant to the analysis of Chapman and Miller's findings:

(1) PRECEDENCE: The subject precedes the object.
(2) PROMINENCE: The subject outranks the object in prominence (here, animacy).

In English (in contrast to, e.g., Japanese), PRECEDENCE is a very strong constraint which results in the ungrammaticality of an OVS order. We hypothesize that the young children in Chapman and Miller's experiment do have the two relevant constraints in their grammar, but have not acquired the right ranking yet.

To test children's comprehension, Chapman and Miller used the four types of sentences illustrated in (3). For each of these sentences, in principle two interpretations are possible, a subject-before-object (SO) and an object-before-subject (OS) interpretation. Adults invariably arrive at the SO interpretation, irrespective of the animacy properties of the arguments, thereby providing evidence for the strict dominance of PRECEDENCE over PROMINENCE in English.

| | | | |
|---|---|---|---|
| (3) | a. | The boy is hitting the girl. (+animate; +animate) | SO – OS |
| | b. | The car is hitting the boy. (-animate; +animate) | SO – OS |
| | c. | The girl is hitting the car. (+animate; -animate) | SO – OS |
| | d. | The car is hitting the boat. (-animate; -animate) | SO – OS |

For young children, the constraint ranking is not so clear yet. For them, animacy might be a more important cue than word order when comprehending a transitive sentence. Following de Hoop and Lamers (to appear), we assume that a combination of two animate or two inanimate nouns violates PROMINENCE once, since the subject does not outrank the object in animacy. PROMINENCE is violated twice if the object outranks the subject (that is, if the object is inanimate and the subject is animate). Finally, PROMINENCE is satisfied if the subject is animate and the object is inanimate. When children have PROMINENCE high-ranked, we therefore expect them to perform best on comprehending sentences with an animate subject and an inanimate object, and to experience most problems when it is the other way around. Chapman and Miller indeed found this pattern in comprehension, with children performing best (93.8%) on sentences like (3c), intermediate (66.5% and 65.2%, respectively) on sentences like (3a) and (3d), and worst (50.1%) on sentences like (3b). The pattern can be illustrated in an OT semantic tableau (Hendriks and de Hoop 2001), where the input is a form as in (3a-d) and the output is a meaning (SO or OS):

(4) OT semantic tableau: from form to meaning

| Input: form | | Output: meaning | PROMINENCE | PRECEDENCE |
|---|---|---|---|---|
| (+anim; +anim) | ☞ | SO interpretation | * | |
| | | OS | * | * |
| (-anim; +anim) | | SO | ** | |
| | ☞ | OS | | * |
| (+anim; -anim) | ☞ | SO | | |
| | | OS | ** | * |
| (-anim; -anim) | ☞ | SO | * | |
| | | OS | * | * |

This tableau predicts that children with the wrong ranking (PROMINENCE >> PRECEDENCE) arrive at the right interpretation of a sentence with an animate subject and an inanimate object, but at the wrong interpretation when the sentence contains an inanimate subject and an animate object. This is basically in accordance with the percentages found by Chapman and Miller. Their interpretations of sentences with two animate or two inanimate arguments are slightly more difficult to explain. In these cases, we expect the SO interpretation to emerge as the optimal interpretation because PROMINENCE cannot make a difference here. But in around 35% of the cases the children get a OS reading. These deviations may be due to the fact that PROMINENCE is actually about discourse prominence rather than about animacy per se. Other factors may influence the perceived prominence relation between the two arguments, such as

visual cues and world knowledge (cf. McClellan et al. 1986). Furthermore, children may initially attend to one or two factors only, rather than to the entire set of factors involved in determining prominence. For these reasons, we expect children to show some variation. Yet, the general pattern seems clear and we can analyse the basic findings in terms of the two constraints and children's reversed ranking of the two.

Recall that our grammar distinguishes the hearer's task from the speaker's task. Which knowledge do English children need to produce the correct word order for a transitive sentence? We claim that production can be modelled by the same grammar as comprehension, i.e., by the same two constraints under the same ranking:

(5) OT syntactic tableau: from meaning to form

| Input: meaning | | Output: form | PROMINENCE | PRECEDENCE |
|---|---|---|---|---|
| (+anim; +anim) | ☞ | SO word order | * | |
| | | OS | * | * |
| (-anim; +anim) | ☞ | SO | ** | |
| | | OS | ** | * |
| (+anim; -anim) | ☞ | SO | | |
| | | OS | | * |
| (-anim; -anim) | ☞ | SO | * | |
| | | OS | * | * |

In the tableau in (5), the input is the event meaning, with the elements between brackets referring to the order of agent and patient. There are two candidate outputs: SO and OS word order. These two forms are evaluated against the two constraints PROMINENCE and PRECEDENCE. However, PROMINENCE is a vacuous constraint in production. When the child wishes to express a given event with particular animacy properties, PROMINENCE is violated or not, irrespective of the word order. Hence, the only constraint that plays a role in production is the word order constraint PRECEDENCE. Therefore, the SO word order is the optimal output for all inputs under both constraint rankings. This straightforwardly explains why all children, also those who do not have the right constraint ranking yet, perform adult-like in production.

## 3. Evidence from Sentence Processing

We have argued that the interaction of two constraints, PROMINENCE and PRECEDENCE, predicts the observed asymmetry in language acquisition. In this section we provide evidence that PROMINENCE plays a role in adult's comprehension as well. In German, like in English, PRECEDENCE outranks PROMINENCE, but OS word order is not

ungrammatical. Because the constraint CASE (de Hoop and Lamers to appear) outranks PRECEDENCE, the OS interpretation is optimal if the first NP bears accusative case.

Although PROMINENCE is low-ranked in German, evidence from German suggests that a violation of PROMINENCE is associated with a significant effect in brain activity. Consider the following (incomplete) indirect questions, which are *ungrammatical* because they contain two nominative NPs, while one should have been accusative:

(6) *… welcher Bischof … der Priester
… [which bishop]$_{NOM}$ … [the priest]$_{NOM}$
(7) *… welcher Bischof … der Zweig
… [which bishop]$_{NOM}$ … [the twig]$_{NOM}$

Frisch and Schlesewsky (2001) found a clear effect in brain activity (a so-called N400 effect) at the second NP in (6) which was lacking in sentence (7). This indicates that people have more problems with processing sentence (6), which contains two animate NPs, than with sentence (7), which combines an animate and an inanimate NP. Because the two sentences differ in the animacy of the NPs, they differ with respect to PROMINENCE. PROMINENCE can be satisfied in (7) (by taking the animate NP as the subject) but never in (6). A similar N400 effect arises at the position of the second NP in the *grammatical* indirect questions in (8) and (9) (Schlesewsky and Bornkessel 2004):

(8) … welchen Bischof der Priester begleitete
… [which bishop]$_{ACC}$ [the priest]$_{NOM}$ accompanied
(9) … welchen Bischof der Zweig streifte
… [which bishop]$_{ACC}$ [the twig]$_{NOM}$ brushed

In (8), PROMINENCE is violated once because subject and object are equal in animacy and hence the subject does not outrank the object. In (9), on the other hand, PROMINENCE is violated twice because the object in fact outranks the subject in animacy. The N400 for sentence (9) compared to sentence (8) thus corresponds to a worse violation of PROMINENCE under the optimal (OS) interpretation, which satisfies the higher ranked constraint CASE. Thus, although adults are not expected to show any production/comprehension asymmetries because they have established the right ranking of the constraints, we do seem to find effects of violations of PROMINENCE with adults as well, namely in their processing of animacy differences.

## 4. Conclusion

On the basis of evidence from first language acquisition we argued that the non-adult

ranking of two violable constraints, PRECEDENCE and PROMINENCE, accounts for the production/comprehension asymmetry with respect to basic word order displayed by young children. Only under the adult ranking does the adult pattern of non-ambiguous forms arise. That both constraints are still active in the adult's grammar, even when one strictly dominates the other, is supported by evidence from sentence processing. This result may have important consequences for our view of the grammar, since it suggests that the grammar is asymmetrical. When used in one direction only, the same set of constraints may generate ambiguity in comprehension, yet no optionality in production.

**Acknowledgements**

We gratefully acknowledge NWO (grants no. 051.02.070 (Hendriks and de Hoop), 015.001.103 (Hendriks), 220-70-003 and DN 30-609 (de Hoop and Lamers)).

**Bibliography**

Chapman, R.S. and J.F. Miller: 1975, Word order in early two and three word utterances: Does production precede comprehension?, *Journal of Speech and Hearing Research* 18, 346-354.

Frisch, S. and M. Schlesewsky: 2001, The N400 indicates problems of thematic hierarchising, *NeuroReport* 12, 3391–3394.

Hendriks, P. and H. de Hoop: 2001, Optimality theoretic semantics, *Linguistics and Philosophy* 24, 1-32.

de Hoop, H. and M. Lamers: to appear, Incremental distinguishability of subject and object, in L. Kulikov, A. Malchukov, & P. de Swart (eds.), *Case, Valency, and Transitivity*, Amsterdam.

Koster, C.: 1993, *Errors in anaphora acquisition*, Ph.D. thesis, University of Utrecht, Utrecht.

McClellan, J., C. Yewchuk and G. Holdgrafer: 1986, Comprehension and production of word order by 2-year-old children, *Journal of Psycholinguistic Research* 15.2, 97-116.

McDonald, J.: 1987, Assigning linguistic roles: The influence of conflicting cues, *Journal of Memory and Language* 26, 100-117.

Prince, A. and P. Smolensky: 2004, *Optimality Theory: Constraint Interaction in Generative Grammar*, Blackwell.

Schlesewsky, M. and I. Bornkessel: 2004, On incremental interpretation: Degrees of meaning accessed during sentence comprehension, *Lingua* 114, 1213-1234.

# A PRESUPPOSITIONAL ACCOUNT OF INDEXICALS

JULIE HUNTER AND NICHOLAS ASHER

Philosophy
The University of Texas at Austin
jhunter@mail.utexas.edu
nasher@mail.utexas.edu

We argue, contra Kaplan, that indexicals should be understood as generating presuppositions that prefer to be resolved in the global context.

According to Kaplan's theory of pure indexicals (1989), terms like *I*, *here*, *now*, *yesterday*, and *the actual* $\phi$ should always pick out their referents from the context of utterance. Moreover, they are directly referential and should always take "wide scope" with respect to modal operators at least in the sense that if George Bush is the actual President, then necessarily, George Bush is the actual President. However, there is evidence that indexicals do not behave as Kaplan predicted. According to Schlenker (2003), for example, Amharic *I* does not always pick out the speaker in the context of utterance. In our paper, we present and analyze examples like (1a,b) which bolster Schlenker's claim that indexicals do not always take wide scope, but offer a different account as to why this is the case.

Consider the following examples:

(1) a. If John Kerry had won the election, the actual President would have been a democrat.

b. If that car were my actual car [pointing at a red sports car], maybe I'd have a girlfriend.

According to Kaplan, *actual* $\phi$ should always be evaluated at the world in the context of utterance; therefore, *the actual President* in (1a) should pick out the President in the actual world, i.e. George W. Bush. However, surveys of native speakers show that the preferred reading of (1a) is one in which *the actual President* picks out John Kerry. In this case, *actual* is evaluated at a world in a counterfactual context introduced by the antecedent of the counterfactuals, contrary to Kaplan's predictions.

We can force *actual* to be evaluated at worlds other than the one supplied by the context of utterance by loading counterfactuals with information that leads to a contradiction, or at least a very unintuitive reading of the counterfactual, if we take *actual* to force the material in its scope to be evaluated at the world provided by the context of use. For example, the antecedent of (1a) asks us to entertain the possibility that John Kerry has won the election. Given this priming, it is more natural to take

*the actual President* to pick out John Kerry—a known democrat—than to interpret the counterfactual as asserting that Bush would have changed political parties had he lost the election. Similarly, if we take *actual* in (1b) to force *car* to be evaluated at the world provided by the context of use, the result is an unnatural identity claim and not the preferred interpretation: 'If I owned that car, maybe I'd have a girlfriend'.

Unlike standard, Kaplan-style treatments according to which *actual* is a modal operator that shifts the world of evaluation to that supplied by the context, we propose that *actual* should be understood as introducing an operator that is relevant to the resolution of presuppositions. In our examples this operator interacts with the presuppositions introduced by a definite description and provides instructions for how the presupposition should be resolved (either bound or accommodated). We follow the general treatment of presuppositions in DRT by van der Sandt (1992) although we disagree with him on one major point. We think that moving away from a modal operator approach to *actual* is important; any modal analysis of the operator that accounts for the data we present forces counterfactuals, conditionals and other modal operators to shift the context (and hence the world supplied by the context), thus accounting for the shift in the world of evaluation of the material within the scope of *actual*. In other words, such an account leads naturally to the presence of monsters. Treating *actual* as giving instructions on how to resolve presuppositions avoids any appeal to monsters, and indeed allows us to forego a special stage of interpretation for indexicals.

Consider (1a) above ((1b) receives a similar treatment). The definite description itself generates the presupposition, $\exists!x(\text{President}(x))$, and *actual* gives specific instructions on where to bind or accommodate this presupposition. This presupposition needs to be satisfied for there to be any hope of the assertion in the consequent of (1a) being true. Our data shows that *actual* forces the resolution of the presupposition in the global context first, or top-level, outermost DRS. We assume that the outermost DRS contains information on the context of utterance such as a speaker, time, and world. If the presupposition finds an acceptable antecedent (i.e. one which yields a sufficiently natural reading) in this context, it will bind to it. If it does not, but accommodation is possible (i.e., the accommodation yields a consistent and pragmatically plausible reading for the sentence as a whole), we will resolve the presupposition by accommodating in the global context. But if binding or accommodation in the global context is not possible as in (1a), *actual* will start searching for an antecedent in a local context and, again, if it finds one there, it will bind to it. If not, then we will attempt to accommodate there. Thus, Kaplan's intuition about indexicals cashes out here as a preference on how to bind the presupposition.

What forces the attachment of the presupposition to the antecedent of the conditional in (1a)? Informally, it amounts to an attempt to make the counterfactual be as plausible as possible (Asher and Lascarides 2003). There are most likely worlds in which John Kerry won the last Presidential election and Bush is a democrat but those won't be intuitively amongst the closest worlds to ours where the election so turns

out. Since counterfactuals are evaluated in the closest worlds to ours where the antecedent is true, choosing a global binding or accommodation for the presupposition would make the counterfactual false or highly implausible. Thus, a pragmatic principle of *charity* (maximize the chances of truth of the assertion) seems an integral part of presupposition attachment.

The semantic effect of *actual* marks expressions like *the actual* $\phi$ apart from other presupposition bearing expressions and forms our principal emendation of van der Sandt's view: whereas van der Sandt's theory predicts that definites prefer to bind locally, we show that when combined with *actual* they prefer to bind globally. To see this, consider

(2) If someone other than George Bush had won the election, the actual President would have been really really angry.

The presuppositions generated by *the actual President* in (2) should on van der Sandt's theory bind to the antecedent of the conditional. Namely, *the actual President* should bind to whomever would have won the election. However, intuitions run counter to this prediction. We predict that *the actual President* in (2), as it prefers a global attachment and nothing blocks that attachment, is perforce evaluated at the world given by the context of utterance. Further, interviews of native speakers have found that *the* $\phi$ and *the actual* $\phi$ differ in their accommodation tendencies. In (3) the tendency is to understand *the actual winner* as bound to the person other than George Bush who counterfactually wins the election, whereas in (4) intuitions differ as to whether the winner is George Bush or someone else.

(3) If someone other than George Bush had won the election, the actual winner of the election would have been happy. (locally bound)

(4) If someone other than George Bush had won the election, the winner of the election would have been happy. (ambiguous)

We take this as evidence that *actual* affects the preferences for binding or accommodating the presuppositions of definites.

Formally, *actual* introduces an operator $\uparrow$ over material in its scope that affects the resolvability, $\Vdash$, of a presupposition, where this includes the pragmatic constraint discussed above.[1] To define $\Vdash$, we simplify binding to a notion of DRS satisfaction $\models$ and accommodation as incorporation of a presupposition $\phi$ into a DRS $K$ $K[\phi$. The interpretation of a subDRS $K$ depends on assignments to discourse referents declared in superordinate DRSs but free in $K$. Let $K_0, \ldots, K_n$ be a sequence beginning with the global DRS $K_0$ such that $K_0 \geq \ldots \geq K_n$, and $\geq$ is the immediate superordination relation on DRSs. Then, $K_0, \ldots K_{l-1}, [K_l, \ldots K_n] \models \phi$ iff $\phi$ is a dynamic consequence of $K_l, \ldots K_n$ relative to any assignment to free variables occurring in $K_l, \ldots K_n, \phi$ that are declared in $K_0, \ldots K_{l-1}$ and satisfy the conditions in $K_0, \ldots K_{l-1}$. Then:

[1] As well as others of the sort discussed in Asher and Lascarides (2003) and Asher (forthcoming).

- $K_0, \ldots, K_i \Vdash \phi$ iff $\exists j \leq i$ and $\exists l \geq 0$ such that $K_0, \ldots K_{l-1}, [K_l, \ldots K_j] \models \phi$ or for some $k$, $0 \leq k \leq j$, $K_k \rhd \phi$, for $\phi$ a normal DRS or DRS condition.
- $K_0, \ldots, K_i \Vdash \uparrow \phi$ iff there is some $j \leq i$ such that $K_0, \ldots K_j \Vdash \phi$ and there is no $k < j$ such that $K_0, \ldots, K_k \Vdash \phi$

Informally, a presupposition is resolvable in a sequence of contexts just in case some subsequence entails the presupposition or it is accommodated at some element in the sequence. Resolving the presupposition means choosing some witness for the existential quantifier. The clause for $\uparrow \phi$ then forces the binding or accommodation of $\phi$ in the outermost context possible.

Assuming the treatment of definites in Asher and Lascarides (2003) and abstracting away from DRS notation, we provide a compositional derivation of the appropriate presupposition for (1a):

- $\|\text{actual}\| = \lambda P \lambda x \uparrow P(x)$
- $\|\text{actual President}\| = \lambda x(\uparrow \text{President}(x))$
- $\|\text{the actual President}\| = \lambda P(p : \exists! y \exists z \exists R(\uparrow \text{President}(y) \wedge R(y, z) \wedge z =? \wedge R =?); a : P(y))$,
  where $p$ and $a$ label the presupposed and asserted components of the DP.
- Ignoring the presuppositions of *the election*, we get:
  $\|1a\| = \exists e \text{Election}(e) \wedge (\text{Win}(e.jk)) \Box\!\!\rightarrow [p : \exists! y \exists z \exists R(\uparrow \text{President}(y) \wedge R(y, z) \wedge z =? \wedge R =?); a : \text{democrat}(y)]$.

Resolving $R$ to identity and $z$ to John Kerry, we have a presupposition that can be bound to the context given by the antecedent of the counterfactual, though it cannot bind or accommodate in the global context:
$\exists e \text{Election}(e) \wedge (\text{Win}(e.jk) \wedge \exists! y(\text{President}(y) \wedge y = jk)) \Box\!\!\rightarrow \text{democrat}(y)$. Our theory requires a dynamic semantics; otherwise the quantification over $y$ in the antecedent of the conditional wouldn't bind the variable in the consequent, nor would the binding from presupposed to asserted constituents make any sense.

Our theory extends to explain the behavior of other indexicals, such as *now* and *here*. Consider;

(5) [Recounting a trip to an airplane museum] That's when I realized that I was now boarding the very plane my grandfather flew during WWII.

(6) Every time I take him to a new restaurant, he says he's been here before.

*Now* and *here* work just like *actual*. They prefer binding in a global context, but sometimes plausibility and other pragmatic constraints will dictate that these presuppositions are satisfied in a more local context, as in examples (5) and (6).

English *I* and *you* trigger presuppositions with more restrictive resolution conditions (unlike *I* in Amharic or even Serbian). Using our enumeration of DRT contexts,

we can expand our operator language to countenance operators of the form $\Uparrow_{K_0}$, which when prefixed to presupposed material $\phi$ means that $\phi$ must resolve at least in $K_0$. The entry for English *I* then is:

- $\|I\| = \lambda P(p :\Uparrow_{K_0} \exists x(\text{Speaker}(x) \wedge x =?)\ ;\ a : P[x])$

If the global context always contains a discourse referent that is linked to the speaker of the context of utterance as in Zeevat (1999), then our semantics predicts that *I* will behave in the way that Kaplan predicted. *You* works analogously; the presupposition always binds to the global context—viz. to the contextually given addressee.

Fiction provides well known examples of shifting uses of indexicals. We stipulate that fiction shifts the $K_0$ context to a fictional one and the actual context of writing or reading is a superordinate DRS $K_{-1}$. Our lexical entries for *here, now, I* and *you* all predict shifting uses of indexicals in fictional contexts. As an example consider *Now our hero was happy* when used in a fictitious story. *Now* searches for a time in the global context first, but such a binding is not acceptable. Thus, it is forced to look to a local context, namely, the one introduced by the fiction.

Our view entails that different presuppositions may require different resolution strategies depending on their environment or associated presupposition triggers. This view contrasts with the standard view of presuppositions, but is well supported by examples of presupposition triggers like *too*, where accommodation is not possible. We can easily formalize the presuppositions for such expressions in our framework by countenancing presuppositions of the form $B\phi$, where $B$ requires material under its scope to be bound.

Turning to demonstratives, we build on Roberts's (2003) theory which takes demonstratives to be presupposition triggers. Demonstratives presuppose a demonstration (where 'demonstration' is construed broadly as in Roberts's paper). The principal difference between her account and ours is on her account demonstratives don't bind locally. A Google search turned up several cases of modally subordinate local binding of the demonstrative *this*.[2] There are also many examples with the demonstrative *that* where local binding or accommodation occurs as in Jeff King's *Every professor admires most that first book of his.* Demonstratives are of a piece with standard indexicals like *actual, here* and *now*. They prefer a binding in the global context via a demonstration, but when pragmatic constraints like charity or semantic constraints like variable binding dictate, they may bind locally. Thus, they would have their presuppositions $\phi$ prefixed with the operator $\uparrow$.

As our data indicates, definite descriptions by themselves don't seem to have a predetermined resolution strategy, but certain *readings* of them do. For example,

[2] For instance:

(7) Assuming that Plan 2 takes effect in 2009, this would mean that the worker would lose 1.2 percent of their scheduled benefit for each of the 39 years (2009 to 2047, when the worker turns 62) included in their benefit calculation. This would leave the benefit at the point of retirement in 2050 at 63 percent (1.012-39) of the scheduled benefit.

the attributive use of a definite description makes the presupposed content part of the local assertion typically. Thus, we could take $\downarrow$ (the converse operator to $\uparrow$) and assign the presuppositional content $\phi$ of a description read attributively as $\downarrow \phi$. Definites understood referentially might well have their presuppositions prefixed again by the operator $\uparrow$. We believe our account extends also to proper names, which generate presuppositions that prefer a global binding or accommodation and thus have the form. $\uparrow \phi$. Our account then immediately explains the behavior of names in modal contexts observed by Kripke (1972).

Finally, with respect to the adverb *actually*, our account predicts that if the adverb includes within its scope presuppositional material, it should behave as though the presuppositions are prefixed with the operator $\uparrow$. But *actually* often takes non presupposed material within its scope. The introduction of the operator must have some semantic or pragmatic effect; so in such a case we hypothesize that $\uparrow$ has a particular discourse function, such as emphasis or correction. For example, suppose you were talking to someone who said *John Kerry is President*. A corrective use of actually would be: *Actually, George Bush is President.*

Our theory takes a middle road between Kaplan and Schlenker. It clearly differs from Kaplan's as it is a single-stage theory which maintains that indexicals need not bind to a global context. But it is not as complicated as Schlenker's in that we do not import contexts as points of evaluation. We explain the shifting of indexicals by minimally extending existing mechanisms in dynamic semantics that account for the resolution of presuppositions. Our account shows that, contrary to Schlenker, shifting indexicals do not require the introduction of monsters.

## Bibliography

Asher, N.: (forthcoming), Varieties of binding in presupposition, *Texas Linguistics Forum*, to appear in a *Festschrift* for Hans Kamp

Asher, N. and Lascarides, A.: 1998, The semantics and pragmatics of presupposition, *Journal of Semantics* 15, 239–299

Asher, N. and Lascarides, A.: 2003, *Logics of Conversation*, Cambridge University Press, Cambridge

Kaplan, D.: 1989, Demonstratives, in J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan*, pp 481–564, Oxford University Press, New York

Kripke, S.: 1972, *Naming and Necessity*, Harvard University Press, Cambridge, MA

Roberts, C.: 2003, Demonstratives as definites, in K. van Deemter and R. Kibble (eds.), *Information Sharing*, CSLI, Stanford

Schlenker, P.: 2003, A plea for monsters, *Linguistics and Philosophy* 26, 29–120

Van der Sandt, R.: 1992, Presupposition projection as anaphora resolution, *Journal of Semantics* 9(4), 333–377

Zeevat, H.: 1999, Demonstratives in discourse, *Journal of Semantics* 16(4), 279–313

# INDEPENDENCE FRIENDLY LOGIC AS A STRATEGIC GAME

THEO M.V. JANSSEN

ILLC
University of Amsterdam
theo@science.uva.nl

**Abstract** *In this contribution investigate an alternative interpretation for Hintikka's Independence Friendly logic. IF logic is not seen as an extensive game, but as a strategic game. We base our semantics on one assumption: the players are rational: they do not play a strategy if there is a better one. In this semantics signalling is not possible. The semantics gives more adequate results for certain linguistic applications.*

## 1. Introduction

Hintikka has raised attention for the role of independence in the semantics of natural languages (see Hintikka and Sandu 1997, Hintikka 1996). His most well known examples concern branching quantifier sentences like

(1) *Some friend of each townsman and some neighbour of each villager hate each other*

In the intended meaning the friend of the townsman can be chosen independently of the neighbour of the villager. This meaning is often discussed, but we will accept that those sentences have such a meaning. But I think that the intended meaning is not captured by the traditional semantics for Hintikka's Independence Friendly logic (henceforth IF logic).

Hintikka argues that independence is a widespread phenomenon and suggests many other applications: in linguistics e.g. the *de dicto - de re* distinction as in *John seeks a unicorn*, but also applications in other fields, e.g. in quantum mechanics.

The the meaning of (1), using some suggestive abbreviations, is represented by:

(2) $\forall x_1 \exists x_2 \forall x_3 \exists x_{4/x_1,x_2} [T(x_1) \wedge V(x_3) \rightarrow F(x_2, x_1) \wedge N(x_4, x_3) \wedge H(x_2, x_4)]$

The subscript $_{/x_1,x_2}$ indicates that the choice of $x_4$ has to be made independent of $x_1$ and $x_2$; it may depend, however, on $x_3$. In this paper we will see IF logic just as first order logic extended with quantifiers like $\exists x_{4/x_1,x_2}$ and $\forall x_{3/x_2}$, in the literature one finds extensions.

The interpretation of IF logic originally is defined by means of a game between two players, the one aiming at confirming the formula, the other at refuting it. Independence is captured by the number of arguments that the strategies have. Some

authors prefer to interpret it as a game between two teams. In later publications the game has been analyzed as an extensive game. Then independence is captured by indiscernible information sets. It has been argued that the rules for playing are not clear and that one has to make rather unnatural assumptions about the properties of the players (Sevenster 2005, van Benthem 2004 ). For instance, they have to forget decisions made before (even their own), and then to remember them again. Furthermore the obtained semantics has been criticized ( e.g. Janssen 2002).

In this paper we will describe the game as an extensive game. The inspiration came from a paper by Sevenster 2005. He considers a spectrum of possible strategies, en several criteria for making a select from that spectrum. A player may choose only strategies of a certain type, other players may know this information and use it to eliminate strategies from their spectrum, etc. In this contribution we investigate properties of one natural assumption from this spectrum: rationality (weakly dominance in his paper). The resulting semantics coincides for many examples with Hintikka's semantics, but is not equivalent. It will be shown that our semantics is more useful for one of the applications in linguistics.

## 2. New semantics

For the ease of discussion, we assume that the formula consists of a quantifier free part, preceded by a list of quantifiers which may be slashed for some variables. We assume that each quantifier has its own indexed variable, and that they appear in order, first the quantifier for $x_1$, then for $x_2$, etc. These restrictions are not essential for the approach, it can be extended to arbitrary IF-formulas.

With a variable $x_i$ is associated a player $p_i$ who determines the value for that variable. Players who determine the values the existential variables aim at truth value **1**, and players for the universal quantifiers aim at truth value **0**. The strategy of the player for $\exists x_3$ may have as arguments all variables with wider scope; for $\exists x_{4/x_2,x_3}$ the variables $x_2, x_3$ are not allowed as argument. A side remark for experts: we do *not* have implicit slashing.

We base our semantics on one assumption: the players use only *rational* strategies. It probably is the easiest to understand if we start with the opposite. A strategy for player $i$ is *irrational* if there is a strategy for her which is better against at least one combination of strategies of the other players and gives the same result in other cases. A strategy is *rational* if it is not irrational. The case that a player has no rational strategies will be discussed separately. Note that a player may have several rational strategies. The notion of rationality is well known in game theory, Apt 2004 gives an overview with many results.

At the start of the game each player determines which her rational strategies are. We define a sentence to be 'true', if any combination of rational strategies guarantees the outcome **1**.

Some examples are the given below; they are played on the domain **N**.

Consider $\exists x_1 \exists x_2 [x_1 = x_2]$. Player $p_1$ has two strategies: choose $x_1 := 0$ and $x_1 := 1$. It depends on the choice by $p_2$ whether the sentence is evaluated **1** or not. None of $x_1$'s two strategies is better, so both strategies are rational. For player 2 there are 4 strategies: $x_2 := 0, x_2 := 1, x_2 := x_1, x_2 := (\mathit{if}\ x_1 = 1\ \mathit{then}\ 0\ \mathit{else}\ 1)$. Of these is $x_2 := x_1$ is the only which against all other strategies yields **1**, so it is (the only) rational strategy. Since it guarantees that the result will be **1**, the sentence is 'true'.

In $\exists x_1 \exists x_{2/x_1} [x_1 = x_2]$. Both players $p_1$ and $p_2$ have two strategies: choose **0** and **1**. For none of the players there is a strategy that improves all others, both play arbitrarily. For some combinations of strategies the sentence yields **1**, for others it yields **0**. So the sentence is not true.

This last example illustrates that our semantics is not equivalent with Hintikka's semantics. In his approach there is only one player $p_\exists$ who plays $x_1 := 0$ for $x_2 := 0$. The strategy $x_2 := 0$ is a constant strategy, i.e. does not have $x_1$ as argument, and therefore is allowed (in his semantics). One might say that the player has to forget her choice for $\exists x_1$, but remembers her strategy. In approaches with more players, one has to allow that they communicate on their strategies although their choices must be independent.

In my opinion it is counterintuitive that the player(s) have to make independent choices, but never the less have a method to guarantee two identical choices. In Janssen 2002 it is argued that there are many other examples where Hintikka's results are against intuitions on independence.

A problem for our approach is the situation that a player has no rational strategy. Consider $\exists x_1 \exists x_2 [x_2 < x_1]$. The only rational choice for $x_2$ is $x_2 := 0$ (because it is better than e.g. $x_2 := 1$). But $x_1$ has no rational strategies: $x_1 := 1$ is better than $x_1 := 0$, but $x_1 := 2$ is even better $x_1 := 1$ etc. For some choices (viz. $x_1 := 0$) the sentence evaluates to **0**. So the definition of truth has to be adapted somehow for this case. We will not pursue this issue because we are here interested in presenting properties of rational strategies.

## 3. Properties of the semantics

Inn this section an important property of the logic will proven: the semantics can be defined bottom up, i.e. compositionally.

**Definition 3..1** *Two strategies $f$ and $g$ are called equivalent if $f$ in any combination of strategies for the other players yields the same value as $g$.*

**Theorem 3..2** *For any rational strategy of a player for $\varphi$ there is an equivalent strategy that depends only on variables that occur in $\varphi$.*

**Proof.** Without loss of generality, we may restrict our considerations to the situation that the player is $n + 1$, and that $x_1$ is the variable that does not occur in $\varphi$. We

will show that a strategy $f_{n+1}(x_1, \ldots, x_n)$ can be replaced by an equivalent strategy $f^*_{n+1}(x_2, \ldots .x_n)$.

As first step we consider the situation that the values of $x_2, \ldots, x_n$ are fixed, say $a_2, \ldots a_n$ respectively. We call such situations A-situations; they differ only with respect of the value of $x_1$. Let $b$ a value for $x_1$ such that the interpretation of $\varphi$ is as maximal as is possible in A-situations (i.e.for $b$ $\varphi$ is **1** if that is possible at all in A-situations), and otherwise $b$ is arbitrary.

Define now

$$f'_{n+1} = \begin{cases} f_{n+1}(b, x_2, \ldots x_n) & \text{in A-situations,} \\ f_{n+1}(x_1, x_2, \ldots x_n) & \text{otherwise} \end{cases}$$

So $f'_{n+1}$ gives the same value for $x_{n+1}$ in all A-situations. Since $x_1$ does not occur in $\varphi$, application of strategy $f'_{n+1}$ yields the same value of $\varphi$ in all A-situations as $f_{n+1}$ does, no matter what the value of $x_1$ is. So after application of $f'_{n+1}$ the value of $\varphi$ is at least as good as the value obtained by following $f_{n+1}$ (because of the choice of $b$). So $f'_{n+1}$ is at least as good as $f_{n+1}$. Since $f_{n+1}$ was rational, $f'_{n+1}$ cannot be better, hence $f'_{n+1}$ is equivalent with $f_{n+1}$.

Next we apply the same procedure to all other situations (values for $x_2, \ldots x_n$), thus obtaining a strategy $f^*$ that is equivalent with $f$, but does not have $x_1$ as argument.

*Summarizing*. Define first $g(y_2, \ldots, y_n)$ by:

$$g(a_2, \ldots, a_n) = \begin{cases} b, \text{ where } b \in A \text{ such that} \\ \quad [\![\varphi]\!]^{\{x_1:b, x_2:a_2, \ldots x_n:a_n, x_{n+1}:f(b,a_2\ldots,a_n)\}} = 1, \\ \quad \text{if such an element exists,} \\ \text{arbitrary otherwise} \end{cases}$$

Define $f^* = f(g(x_2, \ldots, x_n), x_2, \ldots, x_n)$.

Then $f^*$ is equivalent with $f$, but does not have as argument the variable $x_1$ that does not occur in $\varphi$ .

**End of proof.**

A consequence of this theorem is that empty quantifiers can make no difference. The sentence $\forall x \exists x_{/y}[y = x]$ gets the same truth value as $\forall x \exists z \exists x_{/y} y = x$ (both are *not* true). So our semantics is not equivalent with the semantics of Hodges 1997 where the second sentence gets the value 'true'. More generally: signalling is not possible in our approach.

Another example where information from context is used is the following. First note that $\forall x[\exists y_{/x}[x \neq y]$ is not true (in all approaches): it is impossible to make a choice independent of $x$ for $y$ such that the two are unequal. But consider:

(3) $\forall x[\exists y_{/x}[x \neq y] \vee \exists y_{/x}[x \neq y]]$

This sentence is true in the semantics of Hintikka and Hodges. In the left disjunct the player chooses $y := 0$ and in the right $y := 1$, and for the disjunction she chooses

L if $x = 1$. So surprisingly $\varphi \vee \varphi$ is not always equivalent with $\varphi$. In our approach the formula is not true. In the left disjunct there are several rational strategies, e.g $y := 0$ and $y := 1$, and the same in the right disjunct. Whatever strategy is followed for the $\vee$, no guarantee is obtained that all combinations of rational strategies yields **1**.

A second consequence of the theorem is that an inductive definition of satisfaction seems possible. We give here a inductive definition of the case that the slashed quantifier is the last one.

**Definition 3..3** *Let $\varphi(\bar{x})$ be a formula with possibly $\bar{x}$ as free variables. Then by $\mathcal{A} \models^{+}_{G} \varphi(\bar{x})[v]$ is understood that for any combination of rational strategies $\varphi$ get value* **1** *if for $\bar{x}$ we take $v(\bar{x})$ as value.*

**Theorem 3..4** $\mathcal{A} \models^{+}_{G} \exists y_{/\bar{x}}\, \varphi(\bar{x}, y)[v] \Leftrightarrow$ *there is a function* $f\colon (Fr(\varphi) \setminus \{y\}) \to A$ *such that* $(A) \models \varphi(\bar{x}, f(\bar{x})) \wedge \forall z[\exists u[\varphi(\bar{x}, u)] \to \varphi(z, f(z))][v]$

**Sketch of proof.** ($\Leftarrow$) The first conjunct says that $f$ yields **1** for the choices made earlier for $\bar{x}$ (given by $v$). The second conjunct says that if other values for $x$ would have been chosen (captured by $\forall z$) and if there was then a choice that would make the formula **1**, then also strategy $f$ would yield 1. This means that that $f$ cannot be improved by any other strategy, i.e. it is a rational strategy.
($\Rightarrow$) The first conjunct is obvious, the second conjunct says that for other values $f$ cannot be improved, which is a property of rational strategies
**End of proof.**

As a matter of fact, the truth definition mentioned in the last theorem resembles the one put forward by Janssen 2002, one which is based upon investigations of many examples of independence and in an attempt to formalize the notion 'independent choice' upon intuitions on independence'. The present work can be seen as an argument for such a semantics that is now based upon game theoretical arguments.

## 4. A linguistic application

We return to the linguistic application mentioned in the beginnings. We repeat the sentence, and its representation in the logic:

(4) *Some friend of each townsman and some neighbour of each villager hate each other*

(5) $\forall x_1 \exists x_2 \forall x_3 \exists x_{4/x_1,x_2}[T(x_1) \wedge V(x_3) \to F(x_2, x_1) \wedge N(x_4, x_3) \wedge H(x_2, x_4)]$

Consider now the following situation. Among the friends of the townsmen and two groups are distinguished, viz. male and female ones, and the same among the neighbors of the villagers. Assume now that hating is a relation between all pairs of male friends and male neighbors, and also between female friends and female neighbors,

but not between friends and neighbors of different sexes. In this situation the choices of friends for townsmen and neighbors for villagers have to correspond: in both cases male ones, or female ones. So intuitively the choice for $\exists x_{4/x_1,x_2}$ cannot be made independently of the choice for $\exists x_2$. So in this model sentence (4) should not be true. However, (5) comes out true in Hintikka semantics: coordinate the strategies such such that both yield male friends and neighbors respectively. This shows that the required independence is not captured by his interpretation.

Let us now consider our analysis of (5). Suppose now that in the model under discussion a male friend has been chosen. Then a male neighbor must be chosen as well, say Jacob. And if a female friend had been chosen only the choice for a female neighbor would be winning. But the condition of rationality requires that for female friend the original choice Jacob would be winning as well. That is not the case in the given model, so the formula is not true. For this example our semantics gives the desired result, whereas that is not the case for game theoretical semantics. I that also for other applications (e.g. quantum mechanics) that will be the case.

## 5. Conclusions

We have proposed an alternative interpretation for Hintikka's game theoretical semantics for IF logic. It is game theoretically very natural and it yields better results for the investigated linguistic application.

**Acknowledgements**

I thank Merlijn Sevenster for several discussions on IF; without his inspiring paper Sevenster 2005 this could not have been done.

## Bibliography

Apt: 2004, The many faces of rationalizabilty, unpublished paper

van Benthem, J.: 2004, *The epistemic logic behind IF games*, in R. Auxier editor Hintikka Volume, Library of Living Philosophers.

Hintikka, J.: 1996, *The Principles of Mathematics Revisited*, Cambridge University Press

Hintikka, J. and Sandu, G.: 1997, Game-theoretical semantics, in J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic & Language*, pp 361–410, Elsevier Science

Hodges, W.: 1997, Compositional semantics for a language of imperfect information, *Bulletin of the IGPL* 5 (4), 539–563

Janssen, T. M. V.: 2002, Independent choices and the interpretation of IF-logic, *JoLLI* 11, 367–387

Sevenster, M.: 2005, *A player perspective on IF semantic games*, to appear

# WHEN 'WIDENING' IS TOO NARROW

JACQUES JAYEZ LUCIA TOVENA

ENS-LSH, Lyon – jjayez@ens-lsh.fr
Université Paris 7 – tovena@linguist.jussieu.fr

**Abstract** Current proposals that characterise the widening effect of F(ree) C(hoice) I(tems) as an implicature all require additional stipulations and leave a number of observations unexplained. We propose instead that free choiceness results from ensuring that every member of the restriction is equivalent to every other member with respect to the scope. Whereas this general profile is subject to lexical variations within and across languages, it accounts for the family resemblance of FCI.

## 1. The rise of widening

Recent work in semantics based on Kadmon and Landman's (1993) 'widening' characterizes central properties of FCIs as implicatures (Aloni and van Rooij to app; Chierchia 2005; Kratzer and Shimoyama 2002). The general strategy is to use pragmatic principles to derive distribution and intuitive values of FCIs, such as *any* or *qualunque*/*qualsiasi* in Italian. Some approaches embed a disjunction under a modal ($\Box/\Diamond$) operator and derive, for instance, $\Diamond A \wedge \Diamond B$ from $\Diamond(A \vee B)$. Others (Chierchia) aim at unifying Negative Polarity Items (NPI) and FCI. In spite of their differences these approaches share the idea that FCI and NPI lead to 'stronger' interpretations than their plain indefinite counterparts (*any* vs. *a* or *some*, etc.), because they favour the strongest element(s) in given classes of alternatives.

## 2. The fall of widening

One may point out at least four problems raised by these approaches.[1] First, Kamp (1978) and Zimmermann (2000) have shown that FC implicatures are cancellable, as expected for implicatures, see (1a). But this phenomenon has no counterpart with standard FC items (1b,c,d).

(1) a. You may reach the island by boat or by plane, but I don't remember which

[1] A general, as yet unsolved, question concerns the definition of alternatives, which varies across the different works (e.g., see (Aloni and van Rooij) for a discussion of (Kratzer and Shimoyama 2002)). For space reasons, we do not discuss this point here.

[does not implicate : You may reach the island by boat and you may reach the island by plane]

b. *You may consult any file, but I don't remember which$_{English}$
c. *Tu peux consulter n'importe quel fichier, mais je ne sais plus lequel$_{French}$
d. *Puoi consultare qualsiasi file, ma non mi ricordo più quale$_{Italian}$

Second, universal FC items, found in Scandinavian (Sæbø 2001) or in French (Jayez and Tovena 2005a), are not taken into account because widening makes sense for existential/indefinites. Third, modification of the FC phrase ('subtrigging', after LeGrand 1975) sometimes redeems sentences, see **Any student was excluded* (a) vs. *Any student who cheated was excluded* (b). The infelicity of (a) is attributed to the undue extension of the restriction domain (Carlson (1981) followed by many others), and (a) is taken to mean 'every student in the universe was excluded'. As pointed out in (Jayez and Tovena 2005a), this may work for some carefully chosen sentences but does not explain (i) why (b) could not mean 'every student who cheated in the universe was excluded'[2] and (ii) why there has to be a conceptual connection between the modifier property (e.g., having cheated) and the scope property (being excluded), as shown by the infelicity of ??*By pure chance, any student who had a blue shirt also wore jeans*. Fourth, the derivation of the universal value may create difficulties. Chierchia proposes that *any*–like items are existential indefinites that exploit domain widening to generate implicatures that enrich the semantic content. Suppose we have a maximal domain $D$ that satisfies a certain restriction, and its associated lattice $D_{\subseteq}$ over $\wp$(D). Chierchia reasons that choosing the maximal domain $D$ conveys that no point of $D_{\subseteq}$ is excluded as a possible source of instantiation for the variable, say $x$, introduced by *any*. For a sentence like *I saw any student who wanted to see me*, starting from (a), he derives the implicature (b), where $ST$ is the restriction, $S$ the scope and $\Box_{bel,sp}$ corresponds to the speaker's beliefs (Chierchia 2005, p. 30).

(a) $\Box_{bel,sp}(\exists x \in D(ST(x)\ \&\ S(x)) \rightsquigarrow$

(b) $\forall D' \subset D \Box_{bel,sp}(\exists x \in D'(ST(x)\ \&\ S(x)) \Rightarrow \exists x \in D - D'(ST(x)\ \&\ S(x))))$.

(b) is obtained by standard Gricean reasoning, local (recursive) implicature strengthening proposed by Chierchia and a transition from $\neg\Box_{bel,x}\phi$ to $\Box_{bel,x}\neg\phi$ (the 'epistemic step' of (Sauerland to app) discussed in (Tovena 2000)). Together with (a), (b) entails that the speaker believes that every element of $D$ satisfies the scope, i.e. that she saw every student who wanted to see her.

Consider now a sentence such as *Mary saw at least one student of the group that wanted to see her*. Let $D$ be the set of students of the group who wanted to see Mary. The above derivation would presumably be blocked at the level of the epistemic step. Since the speaker expresses uncertainty about the number of students Mary saw, we cannot assign sufficient knowledge to her. Why can't we get the same effect with the micro discourse *Mary saw any student who wanted to see her,* ??*but probably not John* ? According to Chierchia, *any* demands that the alternatives it triggers be used

[2] Nor why *you may consult any file* should not refer to all the files in the world.

(that they be 'active', in his terms). First, the traditional distinction between cancellable (conversational) and non-cancellable (conventional) implicatures is thereby blurred, because requiring that the alternatives remain active is equivalent to assign to *any* an undefeasible universal meaning. We concede that it is not impossible that this meaning has emerged from the kind of implicature derivation indicated by Chierchia. However the crucial point is that it is no longer open to suspension or cancellation and that, in this respect, it constitutes a core meaning of *any*. Second, it is unclear why the speaker did not choose *every* instead of *any* in those cases. What is the specific contribution of the latter? Why would the speaker bother to trigger implicatures, that cannot be cancelled, instead of delivering a plain semantic instruction? Third, why is a sentence like $^{??}$*You must pick any card of the pack* odd while *You may pick any card of the pack* is not? In both cases, the strong implicature of *any* is satisfied and widening is limited by mentioning a particular pack.

## 3. Free choiceness as free choiceness

### 3.1. The intuition

The title of this section makes it clear that for us the phenomenon of free-choiceness has to do with freedom of choice. Indeed, the general idea behind the analysis presented in (Jayez and Tovena 2005a; Jayez and Tovena 2005b) is that a FCI signals that the subsets of members of the restriction are on equal footing with respect to the property of satisfying or not satisfying the scope, or, in other terms, that, at speech time, we cannot single out a particular subset of members by means of their satisfying or not satisfying the scope. This is apparent in simple examples like *Pick any card* where all the cards are presented as equivalent possible choices. The manifestation of this intuition is more convoluted with strictly universal FCIs such as the French item *tout*, which can be anomalous with imperatives (**Prends toute carte* intended as *Pick every card*). Such sentences are out because the set of cards to be picked is determined at speech time as the restriction set itself.

### 3.2. Irreferential and epistemic FCI

For space reason, we will focus only on the irreferential French FC determiner *n'importe quel* (NQ), partially similar to *any*, and on the epistemic *un* N *quelconque* (UQ), partially similar to the Italian *un* N *qualunque* and the German *irgendein*. NQ is an *irreferential* FCI (Jayez and Tovena 2005b), that is, it is not compatible with an interpretation under which a subset of the restriction satisfies or does not satisfy (2a,b) the scope in the current world.[3] UQ is an *epistemic* indefinite FC determiner. It is not compatible with an interpretation under which the speaker knows that certain members of the restriction satisfy or do not satisfy the scope (2c,d).

[3] Actually, the notion of irreferentiality is more general since it extends to any world (real or imaginary), the speaker is referred to (Jayez and Tovena 2005a) for details.

(2) a. $^{*\backslash ??}$Marie a rencontré n'importe quel diplomate
Mary met FCI diplomat
b. $^{*\backslash ??}$Marie a rencontré n'importe quel diplomate, mais pas mon frère
Mary met FCI diplomat, but not my brother
[context: the speaker's brother is a diplomat]
c. Marie a rencontré un diplomate quelconque
d. $^{*\backslash ??}$Marie a rencontré un diplomate quelconque, (mais pas) mon frère

As shown in (Farkas 2002) a formal rendering of such constraints needs to address the three cases of independent, dependent and modalized variables. For modalized variable, it is important to distinguish between the existence/knowledge of the restriction and the determination of the set of satisfiers. For instance, in (3) the speaker may know which files are allowed. What is required is that he cannot *determine* at speech time which file(s) will be consulted or not consulted.

(3) Tu peux consulter un fichier quelconque
You may consult FCI file

As explained in (Jayez and Tovena 2005a), this means that the restrictions on FC items must take into account *all* the accessible worlds, and, technically that the $\Diamond$-operators must be replaced by corresponding $\Box$-operators on the same set of accessible worlds. If $M$ is a right-associative sequence of modal operators, we note $M^{\Box}$ the sequence obtained from $M$ by replacing each $\Diamond$ operator by its $\Box$-counterpart. (4) is a simplified version of the constraint in (Jayez and Tovena 2005b), extended to cover the case of irreferential items. We assume a DRT representation with modal operators

(4) Let $x$ be the variable introduced by a FCI in a DRS K = $[x : R[x]\ S[x]]$, where $R$ is the restriction and $S$ the scope. Let $M$ be the (possibly null) modal sequence characterizing the possible worlds where K is evaluated and $f_{-X}$ be the assignment function that is undefined for any $x \in X$ and coincides with $f$ otherwise. If K is interpreted with the help of an assignment function $f$, then:
1. If the FCI is UQ, the sentence is appropriate only under an interpretation that does not entail:
$\exists x(\Box_{bel,sp}.M^{\Box}(f_{-\{x\}}(R)\ \&\ f_{-\{x\}}(\neg)(S)))$
2. If the FCI is NQ, the sentence is appropriate only under an interpretation that does not entail:
$M'(\exists x(M''^{\Box}(f_{-\{x\}}(R)\ \&\ f_{-\{x\}}(\neg)(S))))$, where $M'$ is the largest initial subsequence of $M$ ending with an operator having wide scope on NQ and $M = M'.M''$.

(4.1) predicts the anomaly of (2d) where the preferred interpretation entails $\exists x\ (\Box_{bel,sp}$ $(\textit{diplomat}(x)\ \&\ \neg \textit{met}(m,x)))$. More complex examples involving an alternation of modals and quantifiers can also be dealt with. For instance, *Il est possible que chaque étudiant ait à lire un livre quelconque* ('It is possible that each student must read some book or other') is predicted to be fine only if the speaker does not identify any book that one of the students should read. (4.2) accounts for the oddness of $^{??}$*Il est*

*possible que Marie ait rencontré n'importe quel diplomate* ('It is possible that Mary met any diplomat'). *Il est possible que* prevents NQ from taking wide scope, producing the offending form $\Diamond(\exists x\phi)$, which entails that a particular diplomat has been met in at least some epistemic alternative, thus violating irreferentiality (see (Horn 2000, p. 170, ex. 60) for a parallel observation on *any*).

### 3.3. The problem of subtrigging

Elaborating on (Jayez and Tovena 2005a) we propose that subtrigging reflects a dependency between properties.

(5) a. ??Any student of the class was excluded
b. ??Tout étudiant de la classe a été renvoyé
c. ??Qualsiasi allievo della classe è stato espulso
d. Any student of the class who had cheated was excluded
e. Tout étudiant de la classe qui avait triché a été renvoyé
f. Qualsiasi allievo della classe che avesse imbrogliato è stato espulso

Subtrigged episodic sentences like the students examples (5d-f) have a standard implicative structure of the form $\alpha$ = [$x$ : *student*($x$) *cheater*($x$)] $\Rightarrow$ [: *excluded*($x$)]. If there are cheating students and $\alpha$ is true, (4) is violated. However, by emphasizing a dependency between properties, subtrigging contributes an additional logical form $\beta = \Box_r([x$ : *student*($x$) *cheater*($x$)] $\Rightarrow$ [*excluded*($x$)]), where $r$ is some modal base corresponding to a set of rules (physical, deontic, legal, etc.). The modal operator is $\Box_r$, thus it is required that there is no $x$ in the current world ($w$) such that $\Box_r$(*student*($x$) & *cheater*($x$) & ($\neg$)*excluded*($x$)), a constraint which can be met when there are cheating and excluded students in the current world. So, the logical form $\beta$ is compatible with (4). $\beta$ is an implicature favoured by the intuitive relation between the properties of the restriction and the scope. However, the existence of such an implicature is not sufficient. For instance, it might be clear from the context that all the students of the class cheated, but this would not license (5a,b,c). So, we assume that the conventionalized licensing condition requires the lexico-syntactic form to be able to trigger an implicature of dependency. In addition to providing an account of subtrigging, this assumption sheds light on the following puzzle. Dayal (2005) strengthens her previous claim that the behaviour of *any* results from the combination of its modal force with the constraint of *Contextual Vagueness* (CV). CV says that *any* is out when it refers to a contextually fixed or salient set. E.g., in (6), *any* is strange because it either refers to absolutely any student in the universe (modal force, (6a)) or to a fixed set (violation of CV, (6b)). Yet, adding a spatio-temporal restriction does not improve (6a), see (6c), and adding a suitable property does improve (6b), see (6d), although the reference to a fixed set is preserved. One might argue that, in (6d), the set of registered students is not 'contextually fixed' since there is no presupposition that such a set exists, in contrast with the definite description in (6b). However, there is no improvement of (6b) if we suppress the presupposition by using an indefinite, see (6e).

(6) a. $*^{\backslash??}$Mary talked to any student
b. $*^{\backslash??}$Mary talked to any student of the class
c. $*^{\backslash??}$At yesterday's meeting, Mary talked to any student
d. Mary talked to any student who had registered for her course
e. $*^{\backslash??}$At yesterday's meeting, Mary talked to any student of a class

## 4. Conclusion

We have shown that the explanatory power of widening is weak and that the notion can even be misleading. Where does its intuitive attractiveness come from? We conjecture that the 'widening effect' is an implicature that reflects the equivalence between the members of the restriction imposed by FCIs: a possible reason for the speaker emphasizing this equivalence is that she does not put any limit on the choice of a particular member.

## Bibliography

Aloni, M. and van Rooij, R.: to appear, Free-choice items and alternatives, in *Proceedings of KNAW*

Carlson, G. N.: 1981, Distribution of Free-Choice *any*, in *7th CLS*, pp 8–23

Chierchia, G.: 2005, *Broaden your views: implicatures of domain widening and the "logicality" of language*, ms. Università di Milano-Bicocca, June 2005

Dayal, V.: 1998, ANY as inherently modal, *Linguistics and Philosophy* 21, 433–476

Dayal, V.: 2005, The universal choice of free choice *any*, in *Linguistic Variation Yearbook 2005*, to appear

Farkas, D.: 2002, Varieties of indefinites., in *SALT XII*, pp 59–83

Horn, L. R.: 2000, Pick a theory, not just *any* theory, in L. Horn and Y. Kato (eds.), *Negation and Polarity. Syntactic and Semantic Perspectives*, pp 147–192, Oxford U. P., Oxford

Jayez, J. and Tovena, L.: 2005a, Free choiceness and Non-Individuation, *Linguistic and Philosophy* 28, 1–71

Jayez, J. and Tovena, L.: 2005b, *'Epistemic' determiners*, Submitted

Kadmon, N. and Landman, F.: 1993, *Any*, *Linguistics and Philosophy* 16, 353–422

Kratzer, A. and Shimoyama, J.: 2002, Indeterminate pronouns: The view from Japanese, in *Third Tokyo Conference on Psycholinguistics*, pp 1–25

LeGrand, J. E.: 1975, Or *and* Any*: The Semantics and Syntax of Two Logical Operators*, Ph.D. thesis, University of Chicago

Sæbø, K. J.: 2001, Free choice items in Scandinavian, *Linguistics and Philosophy* 24, 737–787

Sauerland, U.: to appear, On embedded implicatures, *Journal of Cognitive Science*

Tovena, L.: 2000, Neg-raising: negation as finite failure?, in J. Hoeksema et al. (ed.), *Perspectives on negation and polarity items*, pp 331–356, John Benjamins

# SCALAR USES OF *ONLY* IN CONDITIONALS

SVETA KRASIKOVA, VENTSISLAV ZHECHEV

Seminar für Sprachwissenschaft
Eberhard-Karls Universität Tübingen
svetlana.krasikova@student.uni-tuebingen.de
vzhechev@sfs.uni-tuebingen.de

We argue that sentences of the kind "You only have to go to the North End to get good cheese" can be ambiguous and employ a scalar version of 'only' on one of their readings. So do the exceptive constructions – the cross-linguistic counterparts of 'only have to' sentences. 'Only' is treated as inducing a 'comparative possibility' scale on propositions. The properties of this scale explain the absence of the prejacent presupposition that is usually associated with 'only'. The sufficiency meaning component is argued to be a pragmatic inference, not a part of the truth conditions.

## 1. Introduction

Adverbial *only* has been recently argued to require a special treatment when occurring in sentences expressing sufficient condition. The following sentence, first discussed in von Fintel and Iatridou 2005, proved to be problematic for the existing analyses of *only*:

(1) To get good cheese you only have to go to the North End.

According to the observation in Bech 1955/57, sentences like (1) are equivalent to:

(2) To get good cheese it suffices to go to the North End.

This suggests that *only* can 'reverse' the relation of necessity, expressed by the embedded *have to*, giving rise to the sufficiency reading.

Another striking fact about (1) and others of its kin is that they do not entail the truth of the prejacent, the propositional complement of *only*. In other words, in uttering (1), we do not convey that the embedded anankastic conditional in (3) is true.

(3) To get good cheese you have to go to the North End.

The truth of the prejacent is elsewhere invariably guaranteed and derived in one way or another from the meaning of *only*.

According to von Fintel and Iatridou 2005's cross-linguistic survey of the morphosyntax of the sufficiency modal construction (SMC), as they call (1), a set of languages, like French, Modern Greek, etc., employs a negative adverb and an exceptive phrase instead of *only*:

(4) Si tu veux du bon fromage, tu n'as qu'aller à North End.
if you want of good cheese you NEG have except go to North End

The goal of this paper is to develop a compositional analysis for "only have to" without introducing a new species of *only* in order to account for the lack of the prejacent entailment/presupposition. We claim that the data in question involve scalar uses of *only* and *except*. By integrating the scalarity into the semantics of the SMC, we explain the polarity facts observed in both variants of the construction. Finally, our analysis predicts that (2) is not equivalent to (1) and (4) but rather a pragmatic inference from them.

## 2. Problems with Previous Analyses

To solve the "prejacent problem" von Fintel and Iatridou 2005 pursue a lexical decomposition alternative, assuming that *only* splits into the negation and *except*, drawing on the parallel to the "ne que" construction in French. By allowing the modal to intervene between the two operators they derive the following truth condition for (1):

(5) In some of the good cheese worlds you don't do anything other than going to the North End.

This truth condition combined with the presupposition (6) in the spirit of Horn 1996 does not entail the prejacent.

(6) In all of the good cheese worlds you do something.

The SMC is thus predicted to express the possibility to achieve the goal expressed by the subordinate clause if the condition in the matrix clause is fulfilled. However, this semantics appears too weak to account for the sentences that appear to involve sufficiency in the logical sense:

(7) For the bomb to explode, you only have to press the button.

The condition in (5) would wrongly predict that (7) is true in a world in which pressing the button does not trigger an explosion.

Another proposal, due to Huitink 2005, is to analyze *only* as a universal modal with reversed order of arguments and to use the notion of modal concord to dispense with the semantic contribution of *have to*. The truth condition she arrives at is:

(8) In all North End worlds you get good cheese.

which renders (1) equivalent to (2). This makes wrong predictions in case there are easier ways for obtaining good cheese than going to the North End. If you can as well get good cheese in the nearest shop, (1) is predicted true contrary to our intuitions. The general problem with the modal analysis is that it fails to capture the fact that the SMC does not only introduce a sufficient condition, but also ranks it as the easiest possible.

## 3. Scalar meaning of SMC

Two major inferences from (1) are associated with the contextually provided effort scale:

- none of the ways of getting good cheese (⟦**gc**⟧) ranked higher on the effort scale than going to the North End (⟦**ne**⟧) are necessary
- none of the ways of getting good cheese ranked lower on the effort scale than ⟦**ne**⟧ are sufficient

### 3.1. The Scale

The effort scale ranks propositions according to the degrees of difficulty they are assigned in the world of evaluation. To define the scale, we suggest that the degree of difficulty of a proposition corresponds to its possibility in the actual world. Thus, the comparative possibility relation from Lewis 1973 is used for ranking:

(9) *p* is more difficult in *w* than *q* iff $q <_w p$ (i.e. *p* is less possible than *q* in *w*)

In the degree talk, p is more difficult than q in w iff $D_w(p) < D_w(q)$, where $D_w$ is a function from propositions to their possibility degrees in w.

This ordering allows us to define the relation of sufficiency/necessity between a degree and a proposition based on the corresponding relations holding between propositions:

(10) $\forall q \in D_{st}, d \in D_d , w \in D_s$ (d is sufficient for q in w) ⇔
($\exists p \in D_{st}$, p is d-possible in w: $sufficient_w$ (p, q)

(11) $\forall q \in D_{st}, d \in D_d , w \in D_s$ (d is necessary for q in w) ⇔
($\exists p \in D_{st}$, p is d-possible in w: $necessary_w$ (p, q))

Further on, we assume that in the scalar context necessity and sufficiency are related in the following way:

(12) $\forall q \in D_{st}, d \in D_d, w \in D_s$ (d is sufficient for q in w) ⇔
($\forall d'$: $d' < d \Rightarrow$ d' is not necessary for q in w)

(13) $\forall q \in D_{st}, d \in D_d, w \in D_s$ (d is sufficient for q in w) ⇒
($\forall d'$: $d' < d \Rightarrow$ d' is sufficient for q in w)

(14) $\forall q \in D_{st}, d \in D_d, w \in D_s$ (d is necessary for q in w) ⇒
($\forall d'$: $d' > d \Rightarrow$ d' is necessary for q in w)

### 3.2. The meaning of Only in SMC

We argue that *only* in the SMC is an exclusive particle that operates on a modal, its complement and the contextually determined set of alternatives to the complement. It introduces an existential presupposition, cf. Horn 1996. Crucially, *only* is a scalar operator that uses a comparative possibility scale to order the alternatives.

(15) $[\![\textbf{only}^S]\!] = \lambda C \in D_{s(st)}.\ \lambda p \in D_{st}.\ \lambda M \in D_{(st)(st)}.$
$\lambda w \in D_s$: $\exists r \in C\ [w \in M(r)].\ \forall r \in C\ [\ p <_w r \rightarrow w \notin M(r)]$,
where C is a contextually determined set of alternatives to p and $<_w$ is a partial order on propositions relating their comparative possibility in w

The logical form for (1) is the following:

(16) $([\![\textbf{only}^S]\!](C)([\![\textbf{ne}]\!]))([\![\textbf{have to}]\!]([\![\textbf{gc}]\!]))$

The predicted truth conditions are in (17) and informally in (18):

(17) "You only have to go to the North End to get good cheese" is defined in w iff
$\exists q \in C$: $[w \in [\![\textbf{have to}]\!]([\![\textbf{gc}]\!])(q)]$
If defined, it is true in w iff $\forall r \in C$: $[\ [\![\textbf{ne}]\!] <_w r \rightarrow w \notin [\![\textbf{have to}]\!]([\![\textbf{gc}]\!])(r)]$

(18) A: You don't have to do anything that is less probable than going to the N.E.
P: There is something that you have to do to get good cheese.

From (17) it follows by assumption (12) that the possibility degree assigned to $[\![\textbf{ne}]\!]$ in w is sufficient for $[\![\textbf{gc}]\!]$, or equivalently, that there is a proposition as possible as $[\![\textbf{ne}]\!]$ that is sufficient for $[\![\textbf{gc}]\!]$. This does not derive the sufficiency of $[\![\textbf{ne}]\!]$ directly. However, we argue that the latter inference is a result of pragmatic strengthening: if the speaker knew $[\![\textbf{ne}]\!]$ is not sufficient, he would choose another alternative with the same degree of possibility to make a relevant statement. So the sufficiency of $[\![\textbf{ne}]\!]$ can be considered a conversational implicature.

### 3.3. Neg+Except

*Except* in "Neg+Except" languages mirrors the semantics of the scalar *only*:

(19) $[\![\textbf{except}^S]\!] = \lambda C \in D_{s(st)}.\ \lambda p \in D_{st}.\ \lambda M \in D_{(st)(st)}.$
$\lambda w \in D_s$: $\exists r \in C\ [w \in M(r)].\ \exists r \in C\ [p <_w r\ \&\ w \in M(r)]$

(20) NEG $([\![\textbf{except}]\!](C)([\![\textbf{ne}]\!])([\![\textbf{have to}]\!]([\![\textbf{gc}]\!])))$

### 3.4. Strengthening by Implicature

To account for the non-sufficiency of easier alternatives, we want to strengthen the meaning by the requirement that any possibility degree greater than the one assigned to $[\![\textbf{ne}]\!]$ is necessary. This condition can be derived as a scalar implicature. Note that the presence of the scale associated with the assumptions in (12) - (14) induces an ordering of informational strength on propositions corresponding to alternative degrees of possibility. Suppose that the elements of C are ordered as in (21). Then we expect the informativity ordering of alternative propositions to be as in (22).

(21) …
you go to the nearest shop = ns
you go to the North End = ne
you go to Italy = it
…

(22) $\forall p \in C$: $[[\textbf{ns}] <_w p \rightarrow w \notin [\textbf{have to}]([\textbf{gc}])(p)] \subseteq$
$\forall p \in C$: $[[\textbf{ne}] <_w p \rightarrow w \notin [\textbf{have to}]([\textbf{gc}])(p)] \subseteq$
$\forall p \in C$: $[[\textbf{it}] <_w p \rightarrow w \notin [\textbf{have to}]([\textbf{gc}])(p)]$

According to the requirement that the stronger statements are negated, we strengthen the truth condition by the following implicature:

(23) $\lambda w.\forall p \in C$: $p <_w [\textbf{ne}] \rightarrow \exists r \in C$: $p <_w r$ & $w \in [\textbf{have to}]([\textbf{gc}])(r)$

## 4. Polarity

The scalar reading based on the possibility scale is not the only one available in "only have to" sentences. It should be possible to understand sentences like (1) without comparing alternatives according to their possibility degrees. However, the "non-scalar" reading is available only if the alternative set can be built in a manner different from the one used for the scalar reading:

(24) You only have to take four eggs to bake the cake.
*only*$^S$: the scale of possibility is deduced from the scale of natural numbers and ranks propositions of the type *you take x eggs*.
Non-scalar *only*: the prominent alternatives are *you take 4 eggs*, *you take 500g of flour*, *you take a cup of milk and 4 eggs*…

In the absence of a natural scale, there is no way to distinguish between the alternatives for the two readings and the purely exclusive reading coincides with the scalar one.

Under negation, the non-scalar reading of *only* is not possible whatever might be the case in a positive sentence:

(25) You don't only have to take four eggs to bake this cake…

a) …you also need a cup of milk.

b) # …you need to take five eggs.

To account for the absence of the scalar reading of *only* under negation and the restriction that *except* can only occur in the scope of negation, we treat *only*$^S$ and *except*$^S$ as a PPI and an NPI respectively, drawing on Condoravdi 2002's analysis of *until*$^p$*/erst*. We give a pragmatic explanation for their polarity sensitivity, in the spirit of Krifka 1995.

Negating (1) or (4) results in the following truth condition:

(26) $\lambda w.\ \exists r \in C\ [[\textbf{ne}] <_w r$ & $w \in [\textbf{have to}]([\textbf{gc}])(r)]$

Taking into account the reversed informativity scale of alternative propositions in (27), we get the implicature in (28).

(27) $\exists p \in C$: $[[\textbf{it}] <_w p$ & $w \in [\textbf{have to}]([\textbf{gc}])(p)] \subseteq$
$\exists p \in C$: $[[\textbf{ne}] <_w p$ & $w \in [\textbf{have to}]([\textbf{gc}])(p)] \subseteq$
$\exists p \in C$: $[[\textbf{ns}] <_w p$ & $w \in [\textbf{have to}]([\textbf{gc}])(p)]$

(28) $\lambda w.\ \forall p \in C: [[\![\textbf{ne}]\!] <_w p: [\neg \exists q \in C\ [p <_w q\ \&\ w \in [\![\textbf{have to}]\!]([\![\textbf{gc}]\!])(q)]]$

It can be proved that, due to the denseness of the scale, (26) and (28) are incompatible. Therefore the negation blocks the scalar reading.

## 5. Conclusion

We have proposed a scalar analysis for the SMC that overcomes the problems of the previous analyses. The oddity of "only have to" sentences in scenarios with easier ways for achieving the goal is explained by a scalar implicature violation. The sufficiency inference is derived as a conversational implicature. Additionally, the properties of the assumed scale can be used to account for the polarity sensitivity of $only^S$ and $except^S$.

**Acknowledgements**

We are very grateful to our colleagues and teachers at the SfS, especially Arnim von Stechow, for the instructive discussions and the support. We also thank the participants of SuB 10 (Berlin, October 2005) for their helpful comments.

**Bibliography**

Bech, G.: 1955/57, *Studien über das deutsche verbum infinitum*, Det Kongelige Danske Akademie av Videnskaberne, København

Condoravdi, C.: 2002, Punctual Until as a Scalar NPI, in S. Inkelas and K. Hanson (eds.), in *To appear in The Nature of the Word*, MIT Press, Cambridge, Mass.

von Fintel, K. and Iatridou, S.: 2005, Anatomy of a Modal, MIT

Horn, L.: 1996, Exclusive Company: Only and the Dynamics of Vertical Inference, in *Journal of Semantics*, vol. 13(1): pp 1-40

Huitink, J.: 2005, Analyzing Anankastic Conditionals and Sufficiency Modals, presented at *ConSOLE XIII*

Krifka, M.: 1995, The Semantics and Pragmatics of Polartity Items, in *Linguistic Analysis*, vol. 25: pp 209-58

Lewis, D.: 1973, *Counterfactuals*, Basil Blackwell, Oxford

# DONALD DUCK IS BACK, AND HE SPEAKS SPANISH

LUISA MARTÍ

CASTL
Universitetet i Tromsø
luisa.marti@hum.uit.no

This paper argues that the solution to the "Donald Duck" problem put forth in Schwarzschild 2002 is, despite its simplicity and attractiveness, not viable. Schwarzschild's solution to the problem involves the idea that the domain restriction of indefinites can be a singleton set. This assumption not only solves the "Donald Duck" problem, it also explains why indefinites can take scope outside of syntactic islands in many languages. I show with data from Spanish, however, that there are indefinites whose wide scope is sensitive to islands. If wide scope readings are analyzed using the singleton-set idea, however, their sensitivity to the syntactic environment in Spanish cannot be explained. This suggests that we should reject this assumption. But if so, we no longer have a general solution to the "Donald Duck" problem. I then suggest that solutions to the problem that exploit the semantics of conditionals are to be preferred.

## 1. The "Donald Duck" problem and two of its solutions

The "Donald Duck" problem is the problem that arises when the restriction of a wide scope indefinite is stranded inside of an *if*-clause. Consider the English example in (1a) and its two possible semantic representations in (1b) and (1c), which use material implication. Notice that in (1b) the restriction of the indefinite noun phrase, *relative of mine*, is interpreted inside the antecedent of the conditional, and that in (1c) that restriction is outside the antecedent of the conditional:

(1) a. If *some relative of mine* died this year, I will inherit a house
b. ∃x [x is a relative of mine & x died this year → I will inherit a house]
c. ∃x x is a relative of mine & [x died this year → I will inherit a house]

Donald Duck is not related to me and, intuitively, if he died this year, it wouldn't follow from (1a) that I inherit a house. However, the mere existence of Donald Duck

makes (1b) true: it is true as long as there exists somebody who is not a relative of mine. These very weak truth-conditions prompted Heim 1982 to move the restriction of the indefinite out of the *if*-clause, and Reinhart 1997 (see also Winter 1997) to use choice functions to interpret indefinites. These approaches propose (1c)[1]: now the individual in question has to be a relative of mine, and Donald Duck no longer counts.

Schwarzschild 2002 insightfully points out, however, that (1c) does not solve the problem. This is because (1c) is true as long as I have a relative who did not die this year, but his/her existence doesn't make (1a) true intuitively. In order to really solve the problem, Schwarzschild argues, we need to assume that the contextual restriction of the indefinite noun phrase is a singleton set. In his solution, indefinites are (unambiguously) existential quantifiers whose contextual domain restriction can be a singleton. Indefinites don't move, and, when their domain is not a singleton, they give rise to narrow-scope readings. When their domain is a singleton, we get the appearance of wide scope (cf. Breheny 2003), with the following truth-conditions for (1a):

(2) [∃x x ∈ C & x is a relative of mine & x died this year] → I will inherit a house; C is a singleton set containing a contextually-salient individual

This is what we want: Donald Duck is not this one contextually-salient relative of mine, and neither is just any relative of mine. Only a unique, contextually-salient relative of mine counts. And s/he has to have died for me to inherit the house.

## 2. Spanish *algunos*

This is an elegant solution to the "Donald Duck" problem. The argument in this paper, however, is that it is not viable. The reason has to do with the fact that Schwarzschild's proposal ties the solution to the problem to island-**in**sensitive wide scope. The crucial point is that the mechanism that achieves wide scope and the one that achieves the required, stronger truth-conditions in (2) are the same: the assumption that the domain restriction of the indefinite can be a singleton set. Schwarzschild predicts that indefinites should be able to take scope outside of syntactic islands, since whether the domain restriction of an indefinite is a singleton or not is not something that is expected

[1] The truth-conditions Reinhart 1997 and Winter 1997 propose are as in (i), which is equivalent to (1c). 'CH' stands for 'choice function'. Choice functions are functions of type <et,e>:

(i) ∃f CH(f) & [f(relative of mine) died this year → I will inherit a house]

to depend on the syntactic environment of the indefinite. This might be a good prediction for English (and perhaps other languages), since English indefinites seem to be able to freely take scope outside of syntactic islands. Importantly, however, the Spanish plural indefinite *algunos* can only outscope certain islands[2]. Wide scope (collective) readings for *algunos* are available outside *if*-clauses ((3a)), *wh*-islands ((3b)) and *when*-clauses ((3c))[3]:

(3) a. [Si *algunos hermanos pequeños míos* se rascan la cabeza], mi madre se preocupa, ✓aunque si se rascan la cabeza mis hermanos Pedro y Antonia, mi madre no se preocupa
'If some younger siblings of mine scratch their heads, my mother worries; though if my siblings Pedro and Antonia scratch their heads, my mother doesn't worry'

b. Si se me olvida [quién se pelea con *algunos niños*], el director de la guardería lo apunta en mi historial
'If I forget who gets into a fight with some children, the kindergarten director writes it down on my sheet'

c. Si [cuando *algunos niños* lloran los profesores no reaccionan], el director de la guardería lo apunta en sus historiales
'If when some children cry the teachers do not react, the kindergarten director writes it down on their sheets'

A few empirical remarks are in order. In (3a) (and (4) and (5b) below) I illustrate the procedure to find out about wide scope readings. In (3a), the wide scope reading is compatible with other groups of my younger siblings not causing worry in my mother when they scratch their heads, whereas its narrow scope reading is not. A felicitous continuation that explicitly makes reference to this possibility is taken to indicate that the wide scope reading is available. On a different note, notice that (3b) and (3c) embed the islands in question (*wh*-islands and *when*-clauses, respectively) inside of *if*-clauses. The reason for doing this is that a second scope-bearing element is needed in order to evaluate the scope of *algunos* in these sentences, and *if*-clauses easily provide that second scopal element. One might worry that there is some special property of *if*-

[2] Alonso-Ovalle and Menéndez-Benito 2002 show that the singular indefinite *algún* is also island-sensitive, but they don't look at as many islands as I do here or draw the conclusions I draw.
[3] Compare the behavior of *algunos* in these examples with that of *unos*, another Spanish plural indefinite: *unos* never allows wide scope readings.

clauses that facilitates wide scope readings, but this cannot be, given facts such as (5b) and (5c) below. Notice finally that intermediate readings with *algunos* outside of e.g., *if*-clauses are possible, as shown in (4). (4) suggests that the readings that worry us here are truly scopal (not referential, specific, etc.):

(4) Todos los profesores se enfadan si *algunos alumnos (suyos)* copian en el examen; ✓ y, mira tú por donde, si son Juan y Pedro, el profesor de literatura no se enfada
'Every teacher gets angry if some student (of his) cheats in the exam; and, it is very curious, if it is Juan and Pedro, the literature teacher doesn't get angry'

Wide scope (collective) readings are not available outside relative clauses ((5a)), coordinations ((5b)) or complex NPs ((5c)):

(5) a. Juan aprobó a todos los alumnos [que se leyeron *algunas novelas de Cela*]
'Juan passed every student who read some novels by Cela'

b. Si [mi hermana mayor y *algunos hermanos pequeños míos*] se rascan la cabeza, mi madre se preocupa, # aunque si se rascan la cabeza mi hermana mayor y mis hermanos Pedro y Antonia, mi madre no se preocupa
'If my older sister and some younger siblings of mine scratch their head, my mother worries; though if my older sister and my siblings Pedro and Antonia scratch their heads, my mother doesn't worry'

c. Me preocuparé [si los médicos empiezan a considerar la posibilidad de sedar a *algunos pacientes*]
'I will worry if doctors start to consider the possibility of sedating some patients'

Notice in (5b) that a continuation that explicitly refers to a situation that zooms in on the wide scope reading is impossible. From this I draw the conclusion that the example does not give rise to a wide scope reading.

The wide scope readings of *algunos* in the examples in (3)/(4) are of the stronger kind, like that of our initial (1a) (recall (2)). The strong, wide scope readings Schwarzschild would achieve by making the domain restriction of the indefinite in these examples be a singleton set (with a single plural individual). The crucial question here is: why can't the domain of *algunos* in the examples in (5) also be a singleton set? That would give the indefinite scope outside the islands there, which we don't want. We would want to be able to say that the syntactic environment can have an influence

on whether *algunos* can take wide scope or not, as we would be able to say if its wide scope arose via syntactic movement. But why would the singletonness of the domain restriction have anything to do with the syntactic environment of the indefinite?

Trying to save the singleton indefinite hypothesis by saying that *algunos* is not a singleton indefinite is not helpful: if singletonness is the solution to the "Donald Duck" problem, we would want it to apply to *algunos*. Crucially, the wide scope readings of examples such as (3a) are not true as long as some group of non-scratching siblings of mine exists. Another attempt that does not work is to assume that indefinites move *and* that their domain can be a singleton, i.e., that island sensitivity results from restrictions on movement and the "Donald Duck" problem is solved by *algunos* being a singleton indefinite. This allows problematic derivations in which *algunos* scopes out by movement but its domain is not a singleton, which takes us back to the Heim version of the problem (recall (1c) and that Heim obtained it by moving the restriction of the indefinite outside of the *if*-clause).

## 3. A third solution to the "Donald Duck" problem: Haida 2003

One thing that the two approaches to the "Donald Duck" problem discussed above have in common is that they both assume that conditionals are interpreted as material implication. An alternative approach to the problem, then, is to investigate a different semantics for conditionals. Perhaps the "Donald Duck" problem arises because of the semantics assumed for conditionals and is independent of the semantics of indefinites. This is indeed the approach developed in Haida 2003 to solve his generalized version of the "Donald Duck" problem. Haida assumes a three-valued logic and a semantics for conditionals that makes use of this logic. In this logic, "If $S_1$ then $S_2$" comes out undefined in the crucial case in which $S_1$ is false and $S_2$ is true. This case was at the heart of the matter in our discussion in §1: with material implication in e.g. (1c), the existence of a relative of mine who did not die this year makes the antecedent of the conditional false; this automatically makes the conditional true and is at odds with our intuitions about the sentence. With the new semantics for conditionals, however, we account for the fact that (1a) does not entitle me to a house in case I have a relative who did not die this year.

## 4. Conclusion

The facts about Spanish *algunos* that I have presented here are problematic for many

theories of indefinites, since these theories are usually designed to make indefinites island-insensitive (cf. e.g. the choice function approach in Reinhart 1997 and Winter 1997). However, they have dramatic relevance for Schwarzschild's proposal, since they imply that his elegant solution to the "Donald Duck" problem is not viable. This, we saw, was because Schwarzschild proposes that the same mechanism, singletonness, is responsible for indefinite wide scope and for preventing the "Donald Duck" problem. Hence, solutions to the problem that do not tie it to scope are called for. One such solution is Haida 2003.

**Acknowledgements**

Thanks to Klaus Abels and Yael Sharvit for comments and discussion. All errors are of course my own.

**Bibliography**

Alonso-Ovalle, L. and Menéndez-Benito, P.: 2002, Indefinites and Quantification: the Case of Spanish, handout of a talk presented at *Sinn und Bedeutung VII*, University of Konstanz

Breheny, R.: 2003, Exceptional-Scope Indefinites and Domain Restriction, *Proceedings of Sinn und Bedeutung VII*, pp. 38-53

Gutiérrez-Rexach, J.: 2001, The Semantics of Spanish plural existential determiners and the dynamics of judgment types, *Probus* 13, 113-154

Haida, A.: 2003, An Argument against the Choice-Function Interpretation of *Wh*-in-situ, ms., ZAS, Berlin.

Heim, I.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. thesis, University of Massachusetts, Amherst, MA

Reinhart, T.: 1997, Quantifier Scope: How Labor is Divided between QR and Choice Functions, *Linguistics and Philosophy*, 20, 335-397

Schwarzschild, R.: 2002, Singleton Indefinites, *Journal of Semantics*, 19, 289-314

Winter, Y.: 1997, Choice Functions and the Scopal Semantics of Indefinites, *Linguistics and Philosophy*, 20, 399-467

# A COMPOSITIONAL SEMANTICS FOR LOCATIVES

CÉCILE MEIER

Frankfurt University
C.Meier@lingua.uni-frankfurt.de

Spatial expressions are usually interpreted as relations between two individuals (see e.g. Bierwisch 1996). I am going to argue that a smoother picture of the semantics of locatives arises if we assume that locative prepositional phrases refer to places. This move necessitates the introduction of variables for places into the formal language. I am going to show in detail that an analogous linguistic system underlies the reference to places and the reference to times. Locative modifiers may play the role of frame-setters restricting the reference place. They may set the place of the utterance or the place of an event, state or individual (in analogy to the so-called event time and speech time). Furthermore, expressions like **everywhere** and **nowhere** act as locative quantifiers. In my view, the overall architecture of locative semantics mirrors the properties of other quantificational domains and this view fits nicely into the program of *ontological symmetry* that Schlenker Schlenker 2005 has recently argued for.

## 1. Introduction

The compositionality of spatial expressions has not received too much attention in previous linguistic work. Most papers on locatives concentrate on the conceptual contribution of these expressions (see e.g. Tenbrink 2005 for a valuable survey). And, if they discuss the compositionality of locatives, the authors either only develop a semantics for the predicative use (see e.g. Winter and Zwarts 2000) or only for the modifier use (see e.g. Maienborn 2001) but there is no unified account for all types of grammatical functions, i.e. an account that includes the attributive use and the use as plain arguments.

Maienborn concludes on the basis of empirical data that there are three types of locative modifiers: so-called internal (or V-) and external (or VP-)modifiers and frame setters (or CP-modifiers). Syncategorematically, she introduces three rules of interpretation (named combinatorial templates). And she employs an idiosyncratic variant of variable assignment in order to derive the desired interpretations. Her account is problematic in several respects. In general, all three types of Maienborn's locative modifiers are introduced by the same prepositions in many languages. I know of no language that expresses the combinatorial templates explicitly. By usual

linguistic standards there is a generalization missing here. Why do we need the different interpretation mechanisms? A unified mechanism is preferable.

Life-time effects: It seems that so-called existence-dependent predicates localize an argument of the construction whereas existence independent predicates don't (see Musan 1997 for the notion of existence dependence with respect to temporal properties of individuals). In (1-a) Krause has to be in Spain at the time of his unhappiness in the past. In (1-b) Krause doesn't even have to be alive at the time of his fame in Spain (Examples from Klein 1991).

(1) a. Krause was unhappy in Spain.
b. Krause was popular in Spain.

Maienborn's account fails to explain this life-time effect, familiar from temporal semantics. I will argue that it depends on the semantics of the predicate and not the modifier and conceptual reasoning whether the subjects are localized in the domain of the spatial expression.

Event localization vs. individual localization: In example (2-a) the event of singing is localized in the bathtub, in (2-b) it is only the dog that is there.

(2) a. Jim sang the song in the bathtub.
b. Jim kept the dog in the bathtub.

In order to account for the effect that local adjuncts may have access to participants of a situation, Maienborn introduces a free variable in the logical form of the sentence that rather miraculously ends up having the correct assignment. In my account the correct refernce is part of the semantics of the predicate. (2-a), I am going to interpret as existence dependence of the subject and (2-b) as existence dependence of the object.

Localizing Object Parts: Moreover, some predicates introduce existential quantification over parts of their arguments. (3) is an example for this case. (3) means that there is a part of Jim's face that was red. My account is related to Musan's account of individual time slices but expresses the same information without enriching the ontology with partial individuals.

(3) Jim was red in the face.

Quantification: In Maienborn's account quantification over locations is not an issue. I am going to show that spatial quantificational expressions are best analyzed as relations between two predicates of places. Locative adverbs either restrict the nucleus of such tripartite structures or the restriction, dependent on the information structure of the sentence.

## 2. The Analysis

Ontology: I assume an ontology with individuals, times, worlds and spaces. Whereas times are the real numbers, spaces are triples of real numbers in the so-called three dimensional Euclidean space. Individuals might be located in this absolute space. I am following the tradition in linguistics by introducing an empirical function $\pi$ that assigns an object its space with respect to a world and a time. Herweg and Wunderlich 1991 call this space (i.e. a set of spatial points) the *Eigenspace* occupied by the object in a world at a time; see also Kracht 2004 for a recent mathematically founded version of this concept. Individuals may continuously move over time. With Kracht, I call the space occupied by the object during the time interval the moving takes place its *region*. Regions are subsets of the set of triples of reals, as well. The role of the preposition is conceived as a function that assigns regions to regions (of objects). The preposition **in**, e.g., relates the Eigenspace of an object to the convex closure of the Eigenspace and returns the convex closure minus the Eigenspace of the object; see Kracht 2004. In other words, prepositions assign neighbourhoods to Eigenspaces of objects. They are conceived here as so-called neighbourhood functions; see Herweg and Wunderlich 1991 and Winter and Zwarts 2000 for the concept. Since it is not entirely clear up to now how to formalize the exact contribution to the meaning of the preposition I confine myself to the use of meta language representations for the prepositions.

Formal Language: the semantics uses the types $e$, $t$, $w$, $i$ and $p$ for entities, truth values, worlds, times and spaces. In addition to variables for entities, worlds and time, I introduce variables for spaces into the language. The interpretation function $[\![\ldots]\!]^c$ assigns denotations relative to a context of discourse $c$. In order to interpret complex configurations, I will use functional application, $\lambda$-abstraction and some sort of existential closure.

Locative denotations: Consider in a first step constructions with one-place predicates. In order to account for the life-time effect, I assume that some predicates trigger a presupposition regarding the existence of their argument. **Unhappy** for example depends on the existence of its argument.

(4) $[\![\textbf{unhappy}]\!]^c = \lambda < w, l, t > .\lambda a : a$ exists in $w$ at $t$ in $l$. $x$ is unhappy in $w$ at $t$ in $l$.

Consider the sentence in (5-a). With the semantics for **unhappy** we want to make sure that Krause existed at the time in the past in question and that he was in Spain then and that he was unhappy. Intuitively the evaluation location (the location of Krauses unhappyness) is included in the reference location (the IN-location of Spain). (5-b) states these intuitive truth conditions where @ refers to the actual world.

(5) a. Krause was unhappy in Spain.
b. $\exists t^*[t^* \subseteq \text{PAST1} \;\&\; \exists l[l \subseteq \text{IN+}(\pi(\text{SPAIN}, t^*, @))$
$\&\; \text{UNHAPPY}(@, t^*, l)(\text{KRAUSE})]]$

Tense and aspect are interpreted as definites and semantic relations between properties of time and two times; see von Stechow 2002, for example.

(6) $[\![\text{PAST}_j]\!]^{g,c}$ = is defined if $g(j)$ precedes the speech time $t_c$. If defined, $[\![\text{PAST}_j]\!]^{g,c} = g(j)$.

(7) $[\![\text{PERFECTIVE}]\!]^c = \lambda P.\lambda t.\exists t'[t' \subseteq t \;\&\; P(t')]$

These elements occupy the Tense Phrase and the Aspect Phrase at the level of logical form. In analogy to the temporal domain, I am going to introduce semantic "aspectual" relations between locations and extra syntactic levels in the derivational tree: the ConP (Containment Phrase) and the LocP (Locative Phrase). The head of the first phrase hosts a predicate that I call SUP. This element introduces the containment relation between the location of the predicate (evaluation location) and the reference location and the location of the utterance respectively. They stand in the superset relation.

(8) $[\![\textbf{SUP}]\!]^c = \lambda l.\lambda l'[l' \subseteq l]$

Locative PPs denote reference locations and are definite. These locatives are situated in LocP that dominates ConP, as represented in (9).

(9) $[_{TP}$ PAST1 $[_{AspP}$ PFV $[_{LocP}$ in Spain $[_{ConP}$ SUP $[_{IP}$ Krause [unhappy]]]]]]

The denotation of the locative makes use of the neighbourhood function and the localization function.

(10) $[\![\textbf{in Spain}]\!]^c$ = the IN+-space of Spain in $w$ at $t$
(= IN+$(\pi([\![\textbf{Spain}]\!], t, w))$)

The reading in (5-b) is derived by means of the lexical definitions and existential closure at the level of the Locative Phrase (LocP) and at the level of the Tense Phrase (TP). The denotation of the preposition is constructed by abstracting over the individual, the world and time variable.[1]

Universal quantification is defined as in (11). It relates two locative predicates.

(11) $[\![\textbf{everywhere}]\!]$(L)(P) = 1 gdw. $\forall l[L(l) \;\Rightarrow\; P(l)]$

(12) Krause was unhappy everywhere.

a. $[_{TP}$ PAST1 $[_{AspP}$ PFV [LocP everywhere $[_{LocP}$ C $[_{ConP}$ SUP $[_{IP}$ Krause [unhappy]]]]]]]

b. $\exists t$[PAST1 $\subseteq t \;\&\; \forall l[g(C) \supseteq l \;\Rightarrow\; \text{UNHAPPY}(@, t^*, l)(\text{KRAUSE})]]$

Constructions with predicates that do *not* show the life-time effect are defined as total and not as partial functions.[2]

[1] If the locative remains unarticulated the locative reference must be supplied by a free variable $C$.
[2] Predicates with different orientation trigger different presuppositions on their arguments.

(13) $[\![\textbf{popular}]\!]^c = \lambda < w, l, t > .\lambda a.\ x$ is popular in $w$ at $t$ in $l$.

Let us now turn to predicates of "zero arity" modified with a locative as exemplified in the sentence in (14).

(14) It is warm here.

The locative indexical **here** is conceived as the space that the speaker is talking about in the discourse; see Kratzer 1978 for the different uses of **here**, for example.

(15) $[\![\textbf{here}]\!]^c$ = the space that the speaker has in mind at the speech time $t_c$ in the world of utterance $w_c$.

The zero arity predicate is a function that relates a triple of a world, time and space to a truth value.

(16) $[\![\textbf{warm}]\!]^c$(w,t,l) = 1 iff it is warm in $w$ at $t$ in $l$.

Present tense and imperfective aspect are defined as follows.

(17) $[\![\textbf{NOW}]\!]^c$ = the speech time $t_c$ conceived as a point.

(18) $[\![\textbf{IMPERFECTIVE}]\!]^c = \lambda P.\lambda t.\lambda t'[t \subseteq t'\ \&\ P(t')]$

Intuitively, the sentence expresses the truth conditions in (19). The containment relation seems reversed (compared to constructions with one-place arguments). The location that counts as "here" in the discourse falls within the warm region.

(19) $\exists t[\text{NOW} \subseteq t\ \&\ \exists l[\text{HERE} \subseteq l\ \&\ \text{WARM}(@, l, t)]]$

In order to account for this, I propose the LF representation in (20) for the sentence in (14) with a predicate CAP as in (21) introducing the locational semantic relations.

(20) [$_{TP}$ NOW [$_{AspP}$ IPFV [$_{LocP}$ here [$_{ConP}$ CAP [$_{AP}$ warm]]]]]

(21) $[\![\textbf{CAP}]\!]^c = \lambda l.\lambda l'[l \cap l' \neq \emptyset]$

I have to admit at this point that I do not know what exactly governs the choice of SUP or CAP. But, I believe that it is a question of conceptual knowledge about space. If a locative designates an indivisible space, there is no difference between overlap and inclusion and the reading in (19) could be derived. In temporal semantics a similar fact is usually attributed to the aspectual class of the predicate. In eventives the evaluation time is included in the reference time and in statives it is the other way around. Some constructions (in particular predicative constructions in the present tense) allow for aspectual containment relations in both directions. In these cases, Kamp and Reyle 1993 attribute the differences to the semantic characteristics of the temporal adverb. Some adverbs refer to time spans that are conceived as indivisible (or punctual) and others are conceived to be divisible. In analogy we would have to

classify locations with respect to their divisibility. Whether more fine grained locational structures are available probably depends on experience and conceptualization and the goal of the discourse.

For location independent predicates, I assume that they attribute properties to spatial parts of the arguments.

(22) $[\![\textbf{red}]\!]^c = \lambda < w, t, l > .\lambda a.$There is a part $x$ of $a$ such that $\pi(x, t, w) \subseteq \pi(a, t, w)$ and $x$ is red in $w$ at $t$ in $l$.

**Acknowledgements**

I am grateful for discussions with A. Brasevanu, M. Schwager, A. von Stechow, E. Zimmermann, and the participants of the Semantic Reading Group at Frankfurt University. All shortcomings are nevertheless mine.

**Bibliography**

Bierwisch, M.: 1996, How much space gets into language, in P. B. et.al. (ed.), *Language and Space*, pp 31–76, MIT Press, Cambridge

Herweg, M. and Wunderlich, D.: 1991, Lokale und direktionale, in *Ein internationales Handbuch zeitgenössischer Forschung*, pp 758–785, deGruyter, Berlin

Kamp, H. and Reyle, U.: 1993, *From Discourse to Logic*, Kluwer Academic Publishers, Dordrecht

Klein, W.: 1991, Raumausdrücke, *Linguistische Berichte* 132, 77–144

Kracht, M.: 2004, *Language and Space*, ms, UCLA, Los Angeles

Kratzer, A.: 1978, *Semantik der Rede. Kontexttheorie, Modalwrter, KonditionalsŁtze*, Scriptor, Kronberg

Maienborn, C.: 2001, On the position and interpretation of locative modifiers, *Natural Language Semantics* 9, 191–240

Musan, R.: 1997, Tense predicates and liefe time effects, *Natural Language Semantics* 5, 273–301

Schlenker, P.: 2005, *Ontological Symmetrie in Language: A breif Manifesto*, ms, 2005. To appear in Mind and Langague

Tenbrink, T.: 2005, *Semantics and Application of Spatial Dimensional Terms in Englsih and German*, Technical report, SFB/TR 8 Report No. 004-03/2005, UniversitŁt Bremen

von Stechow, A.: 2002, Temporal prepositional phrases with quantifiers: Some additions to pratt and francez (2001), *Linguistics and Philosophy* 25, 755–800

Winter, Y. and Zwarts, J.: 2000, Vector space semantics: A model theoretic analysis of locative prepositions, *Journal of Logic, Language and Information 9:171-213* 9, 171–213

# COMPARATIVES WITHOUT DEGREES: A NEW APPROACH

FRIEDERIKE MOLTMANN

IHPST, Paris
fmoltmann@univ-paris1.fr

It has become common to analyse comparatives by using degrees, so that *John is happier than Mary* would in fact involve a comparison between the degree(s) to which John is happy and the degree(s) to which Mary is happy. I will argue that such analyses face serious conceptual and empirical problems. I will pursue an alternative analysis on which comparatives involve a comparison between particularized properties or tropes, the kinds of things nominalizations like *John's happiness* refer to.

**Introduction**

In this paper, I will explore a novel analysis of comparatives in which the notion of a concrete property manifestation or trope plays the central role, instead of, as is common, the notion of a degree. The analysis, I argue, has some major conceptual and empirical advantages over a degree-based account. Roughly, on that analysis, instead of analysing (1a) as in (1b) (Cresswell 1976, von Stechow 1984), with the adjective being taken to express a relation between objects and degrees, (1a) is analysed as in (1c) where what is actually compared being the kinds of things nominalizations of the adjective refer to:

(1) a. *John is happier than Mary is.*
b. max d[happy(John, d)] > max d[happy(Mary, d)]
c. [*John's happiness*] < [*Mary's happiness*].

That is, (1a) is taken to mean, roughly 'John's happiness exceeds Mary's happiness'.

There are other versions of the degree-based account, the differences among which won't matter much for our purposes, for example analyses that make use of quantification over degrees (Pinkal 1989, Moltmann 1992) or that make use of measure functions (Kennedy 1999, 2001). What is common to the degree-based accounts is that gradable adjectives express relations between individuals and degrees, or, on Kennedy's account, functions from individuals to degrees.

In a lot of the relevant literature it is left open what degrees actually are, as long as they come with the appropriate ordering. While Cresswell (1976) takes them to be equivalence classes of individuals, they are more often taken to be abstract objects such as numbers or intervals consisting of numbers (see also Klein 1980, 1991). As Kennedy (1999 2001) points out, degrees need to also come with a type or dimension, so that degrees of height are associated with a different dimension than degrees of weight or of beauty (see the next section).

**1. Problems for the degree-based account**

An apparent piece of evidence for the degree-based account is the possibility of overt degree-phrases, expressions which seem to spell out the degree supposedly involved in the meaning of the adjective, as in (2):

(2) a. *John is two meters tall.*
b. *John is two meters taller than Mary.*

The distribution of degree-phrases, however, presents at the same time a serious problem for the degree-based account: degree-phrases are not possible with all adjectives that permit the comparative. The relevant generalization must make a distinction between degree-phrase modifiers of the positive as in (2a) and socalled differential degree-phrases as in (2b), which modify the comparative. The generalization in question seems to be the following (cf. Schwarzschild, to appear). Differential degree phrases in comparatives as in (2a) are possible with any adjective associated with an established measurement scale. Degree phrases with the positive as in (2b), by contrast, are subject to certain general and idiosyncratic lexical restrictions. First, they are impossible with the negative of adjectives (* *two meters small / narrow / short*) (Kennedy 1999, 2001) and with excessives (* *two meters enormous*). Second, degree-phrases are impossible even with adjectives associated with an established measurement scale. Whether such an adjective accepts a degree-phrase seems a matter of lexical particuliarity, differing from language to language (*two kilo heavy* is bad, but the German equivalent *two kilo schwer* is fine) (Schwarzschild, to appear). Thus, a semantic analysis of degree phrases has to account both for the generalizations mentioned above and for the idiosyncratic lexical restrictions imposed by adjectives associated with a measurement scale. Moreover, it needs to account for the fact that the majority of adjectives allowing for the comparative *(beautiful white, soft, strange…)* do not allow for degree phrases even in the comparative construction.

What properties should degrees have? First of all, they are to come with a (total) ordering. For that purpose, they may be conceived of simply as numbers. However, as Kennedy (1999), emphasizes, taking degrees to be numbers is not enough if the aim is also to explain incommensurability in cases of comparative subdeletion, as in (3a), as well as its absence in cases like (3b):

(3) a. # *John taller than Mary is beautiful.*
b. *The table is wider than the sofa is long.*

What is needed in addition to a numerical representation is a type or dimension specifying whether the scale in question is one of height, weight, beauty or whatever. In (3b) the types or dimensions are the same, whereas in (3a) they are different.

One major problem this account faces is: how can it be decided what numbers to choose as the degrees in question? It is quite obvious that there are no good reason to choose any particular system of numbers over any other to represent degrees. Even when a comparative involves an adjective that has a measure system associated with it, there is no reason to choose that system, when no explicit reference is made to it, rather any other

system the speaker may be familiar with. The choice of a measure system thus is, within limits, arbitrary in cases when an established measure system exists. The choice of any particular scale of numbers would be *entirely* arbitrary in cases of adjectives *not* associated measure system (*beautiful, white* etc). The problem of such indeterminacy in the right choice of a system of objects to represent degrees is just the same as is familiar from the philosophy of mathematics, as regards the choice of one set-theoretical construal of numbers over another (Benaceraff 1965). Degrees have certain properties, such as standing in certain relations, being associated with entities (relative to a dimension), but any additional properties such as being '1' as opposed to '100' would be enforced artificially by the choice of a measure system alone. While the problem in the philosophy of mathematics is a fundamental one, in the present case, there is a straightforward way of avoiding it, by using tropes instead of degrees.

A somewhat related problem besides that of the semanticist's choice of appropriate degree objects is for the language *user* to have cognitive access to what he is saying when apparent reference to a degree is made. If a speaker is not able to spell out what degree exactly is involved in the semantic structure of a comparative sentence, how can she actually know what she is saying? This is essentially the meaning-intention problem that Schiffer (1987) discussed in the context of modes of presentations being implicit arguments in attitude reports. Of course, this problem does not arise, if comparatives involve simply quantification over degrees, as in (1a). However, there are cases when reference to particular degrees would take place, as in (4):

(4) *John is as happy as Sue is. Bill isn't that happy.*

In (4), *that* refers to the degree of happiness shared by John and Sue. The meaning-intention problem arising here is that the speaker would not have a clue as to what the degree object is he makes reference to with the utterance of *that* (besides it being the degree of happiness shared by John and Sue)

**2. The proposal**

The point of departure of my analysis is that comparatives such as (5a) are, roughly, equivalent to (5b):

(5) a. *John is happier than Mary.*

b. *John's happiness is greater than / exceeds Mary's.*

(5a) and thus (5b) involves a simple comparison ('greater than' in all cases) between things of the sort John's happiness and Mary's happiness. What are these entities, that is, the referents of nominalizations like *John's happiness*? What *John's happiness* stands for is best viewed as the concrete manifestation of happiness in John, that is, as particularized property or *trope*, to use the term that has become most common in contemporary metaphysics (Williams 1953, Mulligan/Simons/Smith 1984, Woltersdorff 1970, Moltmann 2004). Intuitively, the happiness of John is what you get when you abstract away from all of John's properties except his happiness – which is why tropes have also

been called 'abstract particulars' (Campbell 1990).

Tropes need to be sharply distinguished from related sorts of entities, namely states (which happen to have enjoyed greater popularity in recent linguistic semantics). Tropes are not states in that they are focused entirely on the way the property manifests itself in the individual; states, by contrast, only care about the holding of the property of the individual. This means that tropes, but not states can be compared with respect to the extent to which they instantiate the property in question, as seen below:

(6) a. *John's happiness is greater than Mary's happiness.*

b. ?? *John's being happy is greater than Mary's being happy.*

The approach I am proposing has the following immediate advantages:

[1] No appeal to a dimension or type of degree needs to be made, to provide a natural account and its absence of incommensurability in subdeletion (what matters here will be whether the tropes being compared are of the same nature or not).

[2] No abstract, rarely explicit entities need to be invoked, rather all that is used is entities speakers obviously make reference to independently, namely with nominalizations of adjectives.

[3] The ordering involved can nonetheless be read off the entities directly: given two tropes of redness, for example, it is clear from their nature which one is greater than the other.

For the formal semantic analysis of comparatives using degrees, the idea is that the comparative operator acts also as a nominalization operator, introducing tropes into the semantic structure, just in the way nominalizations do. How then do tropes get introduced, what do they depend on? On might think that tropes depend on just the property expressed by the adjective and the individual. But this is wrong, as this would not capture the fact that tropes constitute the particular *way* the property is manifested in the individual. Thus, in (6a) the two tropes are distinguished not only by one involving Mary and the other John, but also by the fact that one is a greater manifestation of the property than the other. This means tropes depend also on the actual world (as well as perhaps the time in question). Thus (5a) would have to be analysed as in (7a), where f is the same function involved in the semantics of the nominalization as in (7b):

(7) a. f(John, [*happy*], w) > f(Mary, [*happy*], w)

b. $[\textit{happiness}]^w = \{<f(d, [\textit{happy}], w), d> \mid d \in D(w)\}$

Note that adjectives, in the positive, now simply express properties. It is only comparative adjective that will express a relation between individuals and tropes as in (8):

(8) $[\textit{happier}] = \{<d, t> \mid [f(d, [\textit{happy}], w) > t\}$

The formal analysis has another nice feature in that it predicts that, as on Kennedy's (1999) version of the degree-based account, there will be no scope interactions of a degree quantifier with negation and other quantifiers. Thus if (9a) is analysed as in (9b), no degree quantifier can take scope over the negation, and if (10a) is analysed as in (10b),

no degree quantifier can take scope over *few men* either:
(9) a. *John is not happier than Mary*
b. not f(John, [*happy*], w) > f(Mary, [*happy*], w)
(10) a. *few men are happier than Mary*
b. few men x: f(x, [*happy*], w) > f(Mary, [*happy*], w)

**3. Further evidence for tropes from modifiers**

One major piece of evidence for tropes comes from the range of modifiers adjectives may take. Whereas some adjectives allow for modifiers that could be viewed as predicates of degrees modifiers (*very, much, highly*, *two meters, ten kilo*), the full range of modifiers adjectives allow can only be considered predicates of tropes, not of degrees. At least four classes of such modifiers can be distinguished:
[1] modifiers making reference to the particular way the property is manifested, as in *extraordinarily / unusually / exquisitely / strangely beautiful* or *intensely / uniformly / profoundly red*
Clearly degrees cannot be exquisite, strange, intense, or uniform. But tropes, the particular ways properties manifests themselves in objects, naturally can.
[2] modifiers making reference to the perceivability of the property manifestation, as in *visibly / perceivably happy*
Degrees as abstract objects are not perceivable, but tropes, as concrete objects, certainly are.
[3] modifiers making reference to the causal (including emotional) effect of the property manifestation, as in *horribly / astonishingly / fatally weak*
Degrees as abstract objects, on most philosophers' views, are not possible relata of causal relations, but tropes are (for philosophers that accept tropes).
[4] modifiers making reference to the role of the property manifestation as an object of action, as in *deliberately silent*
Degrees as abstract objects certainly are not objects of actions, but tropes quite plausibly can be.

Adjectival modifiers thus generally are best viewed as predicates of tropes. Does this then mean that tropes should be taken as additional arguments of adjectives, parallel to the Davidsonian account of adverbial modifiers, as in the analysis of (11a) in (11b)?
(11) a. *strangely beautiful*
b. λx[∃t[strange(t) & beautiful(x, t))]
I do not think this is required. Rather adverbial modification itself can be viewed as involving a form of implicit nominalization, so that (11a) will be analysed as in (12):
(12) λx[beautiful(x) & strange(f(x, [*beautiful*], w))]

**4. Remaining Issues**

To sum up, the trope-based approach that I have sketched has some major advantages

over the familiar degree-based accounts: first of all in that tropes are independently needed as objects of reference of nominalization; second in that they are far more acceptable philosophically and allow avoiding the problems of indeterminacy and meaning-intention associated with degrees.

Using tropes thus allows for a novel approach to comparatives quite different from the older vagueness-based analyses of Kamp (1975) and Klein (1980) that tried to do without degrees.

The account still has to deal with some important further issues, though, such as the semantics of degree-phrases (in which case I would say reference to degrees does indeed take place), comparatives with adjectives of negative polarity (*John is smaller than Mary*), and the observation that there are in fact two kinds of gradable adjective nominalizations: 'neutral' ones (*John's height*, which allows John to be small) and 'positive ones (*John's tallness*, which requires John to be tall).

## References

Benaceraff, P. (1965): 'What Numbers could not be'. *Philosophical Review* 74, 47-73.
Campbell, K. (1990): *Abstract Particulars*. Blackwell, Oxford.
Cresswell, M. (1976): The Semantics of Degrees'. In B. Partee (ed.): *Montague Grammar*. Academic Press, New York.
Kamp, H. (1975): Two Theories of Adjectives'. In E. Keenan (ed.): *Formal Semantics of Natural Language*. Cambridge UP, Cambridge.
Kennedy, (1999): *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. Garland, New York.
Klein, E. (1980): 'A Semantics for Positive and Comparative Adjectives'. *Linguistics and Philosophy* 4, 1-45.
---------- (1991): 'Comparatives'; In A. von Stechow / D. Wunderlich (eds.): Semantik: Ein Internationales Handbuch der Zeitgenoessischen Forschung. Walter de Gruyter, Berlin.
Moltmann, F. (1992): *Coordination and Comparatives*. MIT, Ph D dissertation.
---------- (2004): 'Properties and Kinds of Tropes: New Linguistic Facts and Old Philosophical Insights'. *Mind* 113, 1-43.
Mulligan, K. / P. Simons / B. Smith (1984): 'Truth makers'. *Philosophy and Phenomenological Research* 44, 287-321.
Pinkal, M. (1989): 'On the logical Structure of Comparatives'. In R. Studer (ed.): *Lecture Notes in Artificial Intelligence*, Berlin, 146-167.
Schiffer, S. (1987): *Remnants of Meaning,* MIT Press, Cambridge (Mass.)
Schwarzschild, R. (to appear): 'Measure Phrases as Modifiers of Adjectives'. *Recherches Linguistiques de Vincennes*.
Von Stechow, A. (1984): 'Comparing Semantic Theories of Comparison'. *Journal of Semantics* 3, 1-77.
Williams, D. C. (1953): 'On the elements of being'. *Review of Metaphysics* 7, 3-18. Reprinted in Mellor/Oliver (eds.), 112-124.
Woltersdorff, N. (1970): *On Universals*. Chicago UP, Chicago.

# SYNONYMY, COMMON KNOWLEDGE, AND THE SOCIAL CONSTRUCTION OF MEANING

REINHARD MUSKENS

Department of Linguistics
Tilburg University
r.a.muskens@uvt.nl

In this paper it is shown how a formal theory of interpretation in Montague's style can be reconciled with a view on meaning as a social construct. We sketch a formal theory in which agents can have their own theory of interpretation and in which groups can have common theories of interpretation. Frege solved the problem how different persons can have access to the same proposition by placing the proposition in a Platonic realm, independent from all language users but accessible to all of them. Here we explore the alternative of letting meaning be socially constructed. The meaning of a sentence is accessible to each member of a linguistic community because the way the sentence is to be interpreted is common knowledge among the members of that community. Misunderstandings can arise when the semantic knowledge of two or more individuals is not completely in sync.

## 1. Introduction

In formal theories of semantics the notion of meaning often seems to be an inherently static descendant from the Platonic world of Forms. As a consequence, semantic relations are predicted to hold once and for all, while divergencies between agents are disallowed. Once, for example, such a theory has established the synonymy of *eye doctor* and *ophtalmologist*, perhaps with the help of a meaning postulate, these expressions must from then on be the same in all contexts and all agents are predicted to believe that John is an ophtalmologist if they believe that he is an eye doctor. It is well-known that such predictions are wrong.

This Platonic view on meaning (inherited from Frege) contrasts with ordinary life where it seems that meaning is something that gets constructed in language communities and between language participants. Is this more pedestrian and earthly perspective compatible with the logical work that has been done so far? And can it contribute to a solution of the well-known problems that the Platonic perspective runs into? In this paper we provide a logical theory of meaning as a social construct that dovetails well with formal semantic theories such as Montague's. We will show how a single agent's *knowledge of meaning* can be formalized and how this leads to a formalization of the *common knowledge about meaning relations* of a set of agents or a linguistic community. This common knowledge is then held to constitute the

social construction of meaning. If it is common knowledge between language participants that *woodchuck* and *groundhog* are synonymous then these words are treated as such and if any of the participants commits himself to the sentence (say) *woodchucks are fertile* it can be concluded that he is committed to *groundhogs are fertile* as well. It can also be common knowledge within a certain group that (say) *woodchuck* denotes the kind of animal that is normally called a woodchuck and this will enable the members of this group to communicate meaningfully about that animal. But the overall theory that we will propose will in itself make no predictions about such form–meaning relations at all and the theory will not exclude the possibility of agents having misconceptions about denotations or relations of synonymy. If your theory of interpretation diverges from mine, I will consider some of your semantic assumptions to be misconceptions and in our conversations miscommunications may arise.

## 2. A Theory of Propositions

Our point of departure will be the Montague-like theory of propositions proposed in Thomason 1980, as streamlined in Muskens 2005. In this theory, each sentence of a given fragment of English is sent to a logical term of a primitive type $p$ (propositions). These logical terms are very close to the syntactic objects they translate. The sentence *Mary is aware that no man talks if a woman walks*, for example, has a translation $(\mathsf{(a\ woman)}\lambda x\mathsf{(mary(aware((if(walk}\,x\mathsf{))((no\ man)talk)))))}$, corresponding to the form of the sentence in which *a woman* has obtained wide scope in some way. Such terms of type $p$ are systematically related to the domain $st$ (sets of possible worlds) with the help of meaning postulates such as the following. [1]

(1) a. $\forall\pi\pi'\tau\tau'(d^0(\pi,\tau)\wedge d^0(\pi',\tau')\rightarrow d^0(\mathsf{if}\ \pi\pi',\lambda i.\tau i\rightarrow\tau' i))$

b. $d^1(\mathsf{man}_{ep}, \mathit{man}_{e(st)})$

Here the $d^n$ ($d$ is for 'determines') are relations connecting objects of type $e^np$ with objects of type $e^n(st)$. With the help of these and other meaning postulates facts like the following are readily established.

(2) $d^0(\mathsf{((a\ woman)}\lambda x\mathsf{(mary(aware((if(walk}\,x\mathsf{))((no\ man)talk)))))},$
$\lambda i.\exists x(\mathit{woman}\ xi\wedge \mathit{aware}\,\mathsf{((if(walk}\,x\mathsf{))((no\ man)talk))}\,\mathit{mary}\ i))$

In this particular case the proposition under consideration is associated with a set of possible worlds, namely those in which it is true that there is a woman such that Mary is aware of the proposition that if that woman walks no man talks. The treatment is hyperfine-grained, for Mary could well be aware of this proposition but not, for example, of its contraposed form. In Muskens 2005 it is shown how terms of type

[1] We simplify here for the sake of exposition, but in (3) below similar meaning postulates are given with the generality that is required.

$$\leftarrow \underline{d(\text{(if((a woman)walk))((no man)talk)}, \tau, j)}$$
$$\text{(3d)} \downarrow \ \tau := \lambda i.\tau_1 i \rightarrow \tau_2 i$$
$$\leftarrow d(\text{((a woman)walk)}, \tau_1, j),\ \underline{d(\text{((no man)talk)}, \tau_2, j)}$$
$$\text{(3g)} \downarrow \ \tau_2 := \lambda i.\neg\exists x[P_1 x i \wedge P_2 x i]$$
$$\leftarrow d(\text{((a woman)walk)}, \tau_1, j),\ \underline{d^1(\text{man}, P_1, j)},\ d^1(\text{talk}, P_2, j)$$
$$\text{(3m)} \downarrow \ P_1 := \textit{man}$$
$$\leftarrow \underline{d(\text{((a woman)walk)}, \tau_1, j)},\ d^1(\text{talk}, P_2, j)$$
$$\text{(3f)} \downarrow \ \tau_1 := \lambda i.\exists y[P_3 y i \wedge P_4 y i]$$
$$\leftarrow \underline{d^1(\text{woman}, P_3, j)},\ d^1(\text{walk}, P_4, j),\ d^1(\text{talk}, P_2, j)$$
$$\text{(3m)} \downarrow \ P_3 := \textit{woman}$$
$$\leftarrow d^1(\text{walk}, P_4, j),\ \underline{d^1(\text{talk}, P_2, j)}$$
$$\text{(3m)} \downarrow \ P_2 := \textit{talk}$$
$$\leftarrow \underline{d^1(\text{walk}, P_4, j)}$$
$$\text{(3m)} \downarrow \ P_4 := \textit{walk}$$
$$\leftarrow$$

Figure 1: A refutation of $\leftarrow d(\text{(if((a woman)walk))((no man)talk)}, \tau, j)$. Selected atoms are underlined. Composition of the substitutions that are found gives the value $\tau = \lambda i.\exists x[\textit{woman}\ xi \wedge \textit{walk}\ xi] \rightarrow \neg\exists x[\textit{man}\ xi \wedge \textit{talk}\ xi]$.

$p$ in fact can function as small programs for computing the truth-conditions of the sentences associated with them. This is in line with the senses-as-algorithms view of Moschovakis 1994. The paper also explains that some of these programs may *diverge*. For example, the programs connected with the Liar and the Truth-teller never halt and no truth-conditions are therefore associated with these sentences. They have a sense but no reference.

## 3. Relativizing to Agents and their Common Beliefs

What is important for present purposes is that in Muskens 2005 most of the real work of the interpretation process takes place on the *object* level of the interpreting logic. It is the meaning postulates that do the work. This allows for the possibility to make the interpretation process dependent upon agents in the following way. First, we make the $d$ relations *world dependent* by providing them with an extra argument for a possible world. E.g. $d^1(\text{woodchuck}_{ep}, \textit{woodchuck}_{e(st)}, j)$ will now mean that the predicate *woodchuck* is determined by the propositional function woodchuck *in world* $j$. In (3) the set of meaning postulates considered in Muskens 2005 is repeated in a slightly generalized form that takes care of the extra world argument that is now added to the $d$ relations. For all notational conventions and for more general explanation the reader is referred to Muskens 2005.

The meaning postulates in (3), in which all free variables are understood to have a universal interpretation, form a *logic program* and therefore combine a declarative interpretation with a procedural one. In Figure 1 a refutation of the *query* $\leftarrow d(\text{(if((a woman)walk))((no man)talk)}, \tau, j)$ is given that simultaneously com-

putes a certain value for $\tau$ ($\lambda i.\exists x[\textit{woman}\ xi \wedge \textit{walk}\ xi] \rightarrow \neg\exists x[\textit{man}\ xi \wedge \textit{talk}\ xi]$). The computation establishes that the proposition (if((a woman)walk))((no man)talk) determines that value in all possible worlds $j$.

(3) a. $d^n(\mathcal{R}, R, j) \rightarrow d^n(\lambda\vec{z}.\textsf{not}\ \mathcal{R}\vec{z}, \lambda\vec{z}\lambda i.\neg R\vec{z}i, j)$
b. $d^n(\mathcal{R}, R, j) \wedge d^n(\mathcal{R}', R', j) \rightarrow d^n(\lambda\vec{z}.\textsf{and}(\mathcal{R}\vec{z})(\mathcal{R}'\vec{z}), \lambda\vec{z}\lambda i.R\vec{z}i \wedge R'\vec{z}i, j)$
c. $d^n(\mathcal{R}, R, j) \wedge d^n(\mathcal{R}', R', j) \rightarrow d^n(\lambda\vec{z}.\textsf{or}(\mathcal{R}\vec{z})(\mathcal{R}'\vec{z}), \lambda\vec{z}\lambda i.R\vec{z}i \vee R'\vec{z}i, j)$
d. $d^n(\mathcal{R}, R, j) \wedge d^n(\mathcal{R}', R', j) \rightarrow d^n(\lambda\vec{z}.\textsf{if}(\mathcal{R}\vec{z})(\mathcal{R}'\vec{z}), \lambda\vec{z}\lambda i.R\vec{z}i \rightarrow R'\vec{z}i, j)$
e. $d^{n+1}(\mathcal{R}, R, j) \wedge d^{n+1}(\mathcal{R}', R', j) \rightarrow$
$d^n(\lambda\vec{z}.\textsf{every}(\mathcal{R}'\vec{z})(\mathcal{R}\vec{z}), \lambda\vec{z}\lambda i\forall x[R'\vec{z}xi \rightarrow R\vec{z}xi], j)$
f. $d^{n+1}(\mathcal{R}, R, j) \wedge d^{n+1}(\mathcal{R}', R', j) \rightarrow$
$d^n(\lambda\vec{z}.\textsf{a}(\mathcal{R}'\vec{z})(\mathcal{R}\vec{z}), \lambda\vec{z}\lambda i.\exists x[R'\vec{z}xi \wedge R\vec{z}xi], j)$
g. $d^{n+1}(\mathcal{R}, R, j) \wedge d^{n+1}(\mathcal{R}', R', j) \rightarrow$
$d^n(\lambda\vec{z}.\textsf{no}(\mathcal{R}'\vec{z})(\mathcal{R}\vec{z}), \lambda\vec{z}\lambda i.\neg\exists x[R'\vec{z}xi \wedge R\vec{z}xi], j)$
h. $d^{n+1}(\mathcal{R}, R, j) \rightarrow d^n(\lambda\vec{z}.\textsf{mary}(\mathcal{R}\vec{z}), \lambda\vec{z}\lambda i.\exists x[x = \textit{mary} \wedge R\vec{z}xi], j)$
i. $d^n(\mathcal{R}, R, j) \rightarrow d^n(\lambda\vec{z}.\textsf{necessarily}(\mathcal{R}\vec{z}), \lambda\vec{z}\lambda i.\forall k[\textit{acc}\ ik \rightarrow R\vec{z}k], j)$
j. $d^n(\mathcal{R}, R, j) \rightarrow d^n(\lambda\vec{z}.\textsf{possibly}\ (\mathcal{R}\vec{z}), \lambda\vec{z}\lambda i.\exists k[\textit{acc}\ ik \wedge R\vec{z}k], j)$
k. $d^{n+2}(\lambda\vec{u}.\textsf{is}\ xy, \lambda\vec{u}\lambda i.x = y, j)$, where $\vec{u}$ contains $x$ and $y$
l. $d^{n+2}(\lambda\vec{u}.\textsf{love}\ xy, \lambda\vec{u}.\textit{love}\ xy, j)$, where $\vec{u}$ contains $x$ and $y$
m. $d^{n+1}(\lambda\vec{v}.\textsf{planet}\ x, \lambda\vec{v}.\textit{planet}\ x, j)$, where $x$ is among the $\vec{v}$
n. $d^{n+1}(\lambda\vec{z}.\textsf{believe}\ (\mathcal{R}\vec{z}), \lambda\vec{z}.\textit{believe}\ (\mathcal{R}\vec{z}), j)$

But there is a second interpretation of these meaning postulates in which the variable $j$ does not range over *all* possible worlds but only over a subset of them, the subset of worlds that is consistent with the semantic assumptions of a certain agent, for example, or the subset of worlds that are in accordance with the common semantic knowledge of a certain community. Let $B$, of type $e(s(st))$, be the *doxastic alternative* relation, so that $B\ john\ ij$ (or $B(john, i, j)$ for readability) formalizes that in world $i$ world $j$ is a doxastic alternative for John.[2] [3] The postulates in (3) can be interpreted with the variable $j$ ranging over John's doxastic alternatives. Technically this can be done by adding $B(john, w_0, j)$ (where $w_0$ is a constant denoting the actual world) as an extra conjunct to the antecedent of all postulates in (3), so that, for example, (3a) becomes $B(john, w_0, j) \wedge d^n(\mathcal{R}, R, j) \rightarrow d^n(\lambda\vec{z}.\textsf{not}\ \mathcal{R}\vec{z}, \lambda\vec{z}\lambda i.\neg R\vec{z}i, j)$. In a computation such as the one in Figure 1 $B(john, w_0, j)$ will now be added to all lines except the first and will act as a constraint on worlds $j$. In fact, the computation in Figure 1 can now be interpreted as *John's* reasoning about the sense-reference

[2] This doxastic alternative relation can be used to render John's *implicit* beliefs; postulate (3n) talks about *explicit* belief.

[3] In the following I will make no distinction between belief and knowledge. While all alternative relations under consideration will be constructed out of agents' *doxastic* alternatives, I will, in accordance with common usage, nevertheless speak of "everyone's *knowledge*" and "common *knowledge*".

relation of a certain sentence, just as the meaning postulates in (3) have been made contingent upon John's implicit beliefs, encoded by John's doxastic alternatives.

Other agents will have their own sets of beliefs and these sets will lead to theories of interpretation that are possibly different from that of John's. If we take any set of possible worlds, there will be a certain set of $d$ relations that hold in every element of that set and a set of possible worlds therefore determines a *theory of interpretation*. If an agent bears the doxastic alternative relation to certain possible worlds, then the theory determined by the set of those worlds may be said to be *that agent's* theory of interpretation. It is also possible to associate a theory of interpretation with *a group of agents*. Let $G$ be such a group ($G$ is supposed to be a constant of type $e(st)$). The relation denoted by $\lambda ij.\exists x(Gxi \wedge Bxij)$ is the alternative relation underlying the modality "everyone in $G$ knows/believes that" (see Fagin et al. 1995). If we take its transitive closure (easily definable within our logic) we arrive at an alternative relation which we will abbreviate as $\mathcal{C}_G$ and which underlies the modality of "common knowledge". The statement $\lambda i.\forall j(\mathcal{C}_G ij \rightarrow \varphi j)$ will be true in all worlds $i$ such that $\varphi$ (of type $st$) holds in all worlds $j$ that are $\mathcal{C}_G$ alternatives to $i$. We abbreviate it as $C_G\varphi$, "it is common knowledge in group $G$ that $\varphi$". For a wealth of information about the common knowledge operator and its logic, see Fagin et al. 1995.

When above we sketched how the meaning postulates in (3) could be relativized to an agent's set of doxastic alternatives, we seemed to be heading for a rather solipsistic notion of meaning, with each agent entitled to his own theory of interpretation and no communication being possible between agents. While this picture may strike some as realistic we take the perhaps overly optimistic view that communication sometimes is possible and this is where the notion of common knowledge comes in. Suppose that the postulates in (3) do not only belong to the meaning postulates that you and I accept but are in fact common knowledge between us. Then I can signal to you that $\lambda i.\exists x[\textit{woman}\ xi \wedge \textit{walk}\ xi] \rightarrow \neg\exists x[\textit{man}\ xi \wedge \textit{talk}\ xi]$ holds in the actual world by getting the proposition (if((a woman)walk))((no man)talk) across. The Fregean assumption of a mysterious realm where propositions reside and where we can grasp them is unnecessary for explaining the possibility of communication. Common knowledge provides a more earthly explanation.[4]

Much of what was said about the sense-reference relation above can also be said about the relation of *synonymy*. A completely fine-grained theory of meaning, such as the ones in Thomason 1980, Moschovakis 1994, or Muskens 2005, will not allow any pair of non-identical expressions to be synonymous. This will evade problems of non-substitutivity but fails to explain in what sense say *woodchuck* and *groundhog* or *ophtalmologist* and *eye doctor* are synonymous. A solution seems to lie in a relativization to the common knowledge of linguistic communi-

[4]Of course, the question how common knowledge can come about or how it can be approximated is a non-trivial one (see Fagin et al. 1995; Vanderschraaf 2002), but in principle it seems to be amenable to rational investigation.

ties. For each $n$, let *syn*$^n$ be a relation of type $(e^np)((e^np)(st))$ with the intended meaning of expressing synonymity between expressions of type $e^np$. For example, *syn*$^1$(woodchuck, groundhog, $j$) is intended to express that *woodchuck* and *groundhog* are synonymous in $j$. It is reasonable to stipulate that $\lambda\mathcal{R}\mathcal{R}'.$*syn*$^n(\mathcal{R},\mathcal{R},j)$ is an equivalence relation for each $n$ and each world $j$ and, moreover, the following interdependency with the $d^n$ relations should hold.

(4) *syn*$^n(\mathcal{R},\mathcal{R},j) \wedge d^n(\mathcal{R},R,j) \rightarrow d^n(\mathcal{R},R,j)$

If *syn*$^1$(woodchuck, groundhog, $j$) now holds of all $j$ such that $\mathcal{C}_Gij$ for some group $G$, the members of that group will have common knowledge that *woodchuck* and *groundhog* are synonymous and denote the same animal. The notion of synonymy has thus been relativized to groups as well and now has a social interpretation. In a future longer paper we hope to investigate some of the consequences of this perspective on synonymy with regard to some classic foundational puzzles of semantics.

## 4. Conclusion

We have sketched a theory in which central notions of semantics are relativised to a group interpretation. This brings formal semantics more in line with certain standard linguistic insights than it was before. The Saussurean insight that the form–meaning relation is *arbitrary* dovetails well with the present set-up. That there may be individual *divergencies* from the form–meaning relation accepted by a certain group is also easily explained, as is the possibility for the form–meaning relation of a certain group to shift over time. The model also strongly suggests that it is advantageous for a group to have a stable and large common theory of interpretation and that it may be advantageous to an individual to adopt that common theory.

## Bibliography

Fagin, R., Halpern, J., Moses, Y., and Vardi, M.: 1995, *Reasoning about Knowledge*, MIT Press, Cambridge, MA

Moschovakis, Y.: 1994, Sense and Denotation as Algorithm and Value, in *Logic Colloquium '90 (Helsinki 1990)*, Vol. 2 of *Lecture Notes in Logic*, pp 210–249, Springer, Berlin

Muskens, R.: 2005, Sense and the Computation of Reference, *Linguistics and Philosophy* 28(4), 473–504

Thomason, R.: 1980, A Model Theory for Propositional Attitudes, *Linguistics and Philosophy* 4, 47–70

Vanderschraaf, P.: Summer 2002, Common Knowledge, in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, `http://plato.stanford.edu/archives/sum2002/entries/common-knowledge/`

# MONOTONE AMAZEMENT

RICK NOUWEN

Utrecht Institute for Linguistics OTS
Utrecht University
rick.nouwen@let.uu.nl

## 1. Evaluative Adverbs

Adverbs like *amazingly*, *surprisingly*, *remarkably*, etc. are derived from so-called evaluative predicates. There is a subtle difference in meaning between (1-a), where the evaluative adverb *amazingly* is used ad-sententially, and (1-b), where *amazingly* seems to modify the determiner *few*.

(1) a. Amazingly, few people came to my party.
b. Amazingly few people came to my party.

Whereas it follows from (1-a) that few people attended my party, this does not necessarily follow from (1-b). It seems that *amazingly few people* are not necessarily *few people*. For instance, it could be that my parties, including this one, are always extremely well-attended. Although this means that many people came to the party, I could still use (1-b) to express that they were by far not as many as I had anticipated.

The meaning contrast is clearer in examples with an absolute gradable adjective. Katz 2005 points out that the entailment in (2-a) holds, whereas the one in (2-b) obviously does not.

(2) a. Surprisingly, the glass is empty. $\Rightarrow$ The glass is empty.
b. The glass is surprisingly empty. $\not\Rightarrow$ The glass is empty.

As Katz points out, however, the contrast turns up with relative gradable adjectives as well. It is possible to consider someone quite short, but at the same time think of this person as *surprisingly tall*, given that, for instance, I had expected him or her to be even shorter. Being *surprisingly tall* therefore does not entail being *tall*.

This shows that at least two occurences of evaluative adverbs should be distinguished. Ad-sentential adverbs as in (1-a) and (2-a) simply comment on a proposition. So, (1-a) says that few people attended my party and that I consider this fact to be amazing. The cases in (1-b) and (2-b), however, are less straightforward. They seem to involve an evaluation of the degree expressed by the adjective or the determiner. One could, for instance, propose an analysis as in (3).

(3) Cody is amazingly tall. = The degree to which Cody is tall is amazing.

Morzycki 2004, however, argues against an all too simplistic paraphrase as in (3). For instance, suppose that Cody is extremely short. In fact, he is so short, that it is amazing. It would now follow from an analysis as in (3) that Cody is *amazingly tall*, simply because the degree to which he is tall is amazing. Clearly this is not right. For another example, suppose it has been established that Cody is exactly as long as his neighbour is, not a millimetre shorter or longer. Consequently, both neighbours could find the degree to which Cody is tall amazing. But that does not commit them to consider Cody *amazingly tall*.

Despite these remarks, I will propose an analysis of structures like *amazingly tall* which is not very different from (3). I will propose a relatively simple semantics of evaluative adverbs that rests on the crucial assumption that evaluative predicates are monotone. The starting point for my investigation is therefore the interaction of evaluative adverbs with the polar orientation of expressions.

## 2. Evaluation and Polarity

Katz 2005 discusses polarity effects with evaluative adverbs. The examples in (4), for instance, can both be uttered about the same pool of water. They differ, however, in what is expressed about the speaker's expectations.

(4) a. The water is surprisingly warm. (the water was expected to be less warm)
b. The water is surprisingly cold. (the pool was expected to be warmer)

Katz proposes to account for these effects by assuming that the semantics of evaluative degree modifiers involves universal quantification over degrees. So (4-a), for instance, is paraphrased as: 'there is a degree $d$ such that the water is $d$-warm and for every degree $d' \geq d$ it is the case that would be surprising that the water is $d'$-warm'. The domain of quantification for the universal quantifier depends on the polarity of the degree predicate that is being modified. For *warm* the relative ordering relation is '$\geq$', but for a negative scalar predicate like *cold* the ordering relation will be '$\leq$'.

This analysis is partly based on the observation that the polarity effects seem to disappear when the evaluative predicate is not used adverbially. The first sentence in (5), for instance, does not say anything specific about what was expected.

(5) It is surprising that the pool is 60% full. We thought it'd be totally full or empty.

Katz concludes from this that the morpheme *-ly* is responsible for effects like those in (4) and that, consequently, it is this morpheme that adds universal quantification to the meaning of the evaluative predicate. I will argue, however, that the effects in (4) are due to inferences which are triggered solely on the basis of the monotonicity of both the evaluative predicate and the degree predicate. Evidence for this comes from the realm of nominal quantifiers. Although the examples in (6) have subtle meaning

differences, they all express a similar disappointment about how many people turned up.

(6) a. Surprisingly, few people turned up. (?less/more were expected)
b. Surprisingly few people turned up. (#less/more were expected)
c. It is surprising that few people turned up. (?less/more were expected)

The example in (6-c), together with the contrast in (7-a) and (7-b), show that the polarity effects are present in all the different guises evaluative predicates take.

(7) a. It is surprising that many people turned up. (less/#more were expected)
b. It is surprising that exactly fifty people turned up. (less/more were expected)

Apparently, evaluative predicates interact with monotonicity. I will therefore propose an analyses for evaluations of quantifiers and quantified statements as well as for evaluations of degree expessions which is based entirely on the assumption that evaluative predicates are monotone. That is, all forms of evaluative predication trigger polarity effects, since all these forms essentially involve a monotone predicate.

Monotonicity is defined as follows. Let $P$ be some operator on propositions.

(8) a. $P$ is MON$\uparrow$ iff $p \rightarrow p' \;\Rightarrow\; P(p) \rightarrow P(p')$
b. $P$ is MON$\downarrow$ iff $p \rightarrow p' \;\Rightarrow P(p') \rightarrow P(p)$.

Viewed as propositional operators, most evaluative predicates seem to be downward entailing. If I find $p'$ *amazing* (or *surprising, remarkable, fortunate, unfortunate, revolting, nauseating,* etc.) then the same would hold for any stronger proposition $p$. For instance, if the fact that Cody came to my party is remarkable, then it would be equally remarkable had Cody come to my party early. Indications of expectation are examples of upward entailing operations. I cannot expect Cody to come to my party early, without expecting him to come to my party.

In my view, these simple monotonicity inferences are responsible for the polarity effects observed in the previous section. It follows from *Many people came to the party* that *Some people came to the party*. So, if some people turn up and this fact surprises me, then I would have been surprised as well in case many people had turned up. This indicates that my expectation was that no students came.

The main advantage of this analysis is that there is a straightforward relation between adjectives like *surprising* and adverbs like *surprisingly*. In contrast to the proposal in Katz 2005, with this account no additional mechanism is needed to explain polarity effects with evaluative adverbs. These effects occur since the sentence in the scope of the evaluative predicate allows for monotonicity inferences.

There is also a straightforward account for (5). Given that *60% full* is (or at least can be) construed as being non-monotone, nothing follows about other propositions being surprising as well. Since (5) does not commit the speaker to being equally surprised about less or more people being in the pool, nothing can be said about

what causes the surprise. (Hence the felicitous open continuation). Similarly, the question marks in (6-a) and (6-c) are due to the potential exhaustification of the quantificational statement, rendering it non-monotone as well.

As I will explain next, taking evaluative predicates to be monotone will clarify the semantics of evaluative adverbs to considerable degree.

## 3. Evaluation and the Standard of Comparison

An important difference between the ad-sentential and the modifier use of evaluative adverbs is that the latter is more restricted. As becomes clear from (10), only gradable adjectives can be combined with evaluatives. In this sense they parallel degree modifiers like *very*. Similar observations can be made with respect to determiners.

(9) a. Surprisingly, Vic is tall/dead/imprisoned.
b. Surprisingly, Vic has many/some/twenty friends.

(10) a. Vic is surprisingly/very tall/#dead/#imprisoned.
b. Vic is surprisingly/very *(well) educated.
c. Vic has surprisingly/very many/#some/#twenty friends.

These observations support the intuition that was discussed in the first section: when someone is said to be *amazingly tall*, it is not the fact that (s)he is tall that is amazing, but rather the degree to which (s)he is tall. Obviously, a paraphrase like this presupposes gradability.

Gradable adjectives have a positive form in which the relevant degree is contextually determined. So, whereas *2 feet tall* denotes the set of entities that are 2 feet tall, $[_{\text{AP}}$ *tall* $]$ denotes the set of entities that are tall to a sufficient degree determined by some contextual norm, the so-called standard of comparison. It is standard practice to assume that this positive form involves some silent operation, referred to below as $\varnothing$ (see Kennedy 2005 and references therein). It is this operation that provides the contextual standard of comparison. I propose that an adverb like *surprisingly* can perform a similar operation. It too provides a standard of comparison, but unlike $\varnothing$, this standard is not contextual, but rather determined by the criterion of causing surprise. So, in parallel to $\varnothing$, the semantics of *surprisingly* is a function from degree predicates to sets of individuals.

(11) a. $\varnothing \rightsquigarrow \lambda P.\lambda x.\exists d[\mathbf{C}(d)\ \&\ P(d)(x)]$
b. *surprisingly* $\rightsquigarrow \lambda P.\lambda x.\exists d[\textsc{surprising}(^\wedge P(d)(x))\ \&\ P(d)(x)]$

The operator $\varnothing$ takes a degree predicate and returns the set of those individuals for which the predicate holds for some degree $d$ which corresponds to the contextual standard for this predicate (where **C** represents the contextual selection).[1]

[1]This is one of many possible proposals in the literature for how the standard of comparison is incorporated in the positive form. It is not crucial to the analysis, however, that I happen to have chosen this

At first sight, it might seem that this analysis runs into the same problems as did the paraphrase in (3), since (12-c) suggests that something is *surprisingly tall* if the degree to which it is tall is surprising.

(12) a. *tall* $\rightsquigarrow \lambda d.\lambda x.\text{TALL}(x) = d$
b. $[_{\text{AP}} \varnothing$ *tall*$] \rightsquigarrow \lambda x.\exists d[\mathbf{C}(d)\ \&\ \text{TALL}(x) = d]$
c. $[_{\text{AP}}$ *suprisingly tall*$] \rightsquigarrow \lambda x.\exists d[\text{SURPRISING}(^{\wedge}\text{TALL}(x, d))\ \&\ \text{TALL}(x, d)]$

All depends, however, on the semantics one assumes for the degree predicate (TALL). One possibility is that degree predicates are taken to be monotone relations (Heim 2000). That is, if something is $d$-tall, then it is also $d'$-tall for any $d' \preccurlyeq d$. Given this assumption, the objections against a simple analysis as in (12-c) disappear. Take the case of the incredibly short Cody who despite his surprising height cannot be called *surprisingly tall*. If I am surprised at $\text{TALL}(\text{c}, d)$ then, given the monotonicity of surprise and the monotonicity of the degree predicate, it follows that I would be equally surprised had Cody been taller. This explains why short Cody cannot be called *surprisingly tall*.

Clearly, what I should say of Cody is that he is *surprisingly short*. Assuming, once again, that SHORT is monotone, it follows that if someone is $d$-short, (s)he is also $d'$-short for any lesser degree of shortness.[2] So, if I am surprised at Cody's degree of shortness, it follows that in case Cody had been even shorter, I would have been surprised as well. Thus, the monotonicity of degree predicates makes *surprisingly short* a suitable and *surprisingly tall* an unsuitable description of Cody.

Let me finaly return to the contrast I observed at the beginning of the paper. In ad-sentential position, an evaluative adverb is factive. So, from [*Surprisingly, S*] it follows that *S*. This is easily explained if we view the adverb as a parenthetical comment on the assertion of *S* (cf. Potts 2005). As a degree modifier, however, the evaluative adverb shows no traces of factivity. In particular, it was noted that being *surprisingly A* does not entail being *A*. In the absence of an evaluative modifier, the adjective is in its positive form. So, someone is tall if and only if one is at least as tall as some contextually determined norm, the standard of comparison. This standard, however, could be based on a lot of things, among which one's expectations, one's desires, one's obligations, etc. Whereas a modifying evaluative adverb fixes the standard to a specific mode of evaluation, the positive form can be interpreted with respect to a different, potentially higher, standard. This explains why, for instance, something can be *surprisingly tall* without being *tall*.

---

particular one. See Kennedy 2005 for extensive discussion.

[2] An elegant and arguably necessary means of relating degree predicates that form polar opposites, like *tall*/*short*, is taking degrees to correspond to intervals (or *extents*). One can then distinguish between positive degrees of the form $\langle 0, d\rangle$, and negative degrees of the form $\langle d, \infty\rangle$. Viewing degrees as (ordered) sets of values, one can moreover define an ordering relation that applies to both negative and positive degrees: $d \succcurlyeq d' \Leftrightarrow d \cap d' = d'$. See Kennedy 2001 and references therein for discussion.

## 4. Conclusion

I have defended a simple analysis of evaluative adverbs modifying gradable adjectives. Crucial to this analysis is the interaction of forms of monotonicity. Based on the hypothesis that evaluative predicates are monotone, I have been able to maintain, first of all, that both evaluative predicates and adverbs mark polar orientation and, second, that when evaluative adverbs modify gradable adjectives, they express evaluations of degrees.

As far as I can see, the analysis extends without problems to cases where evaluatives combine with gradable determiners like *many* and *few*. However, I leave the precise details of the correspondence of vague determiners to degree predicates to further research.

## Acknowledgements

This work was supported by a grant from the Netherlands Organisation for Scientific Research (NWO), which is gratefully acknowledged.

## Bibliography

Heim, I.: 2000, Degree operators and scope, in *Proceedings of SALT X*, CLC Publications, Ithaca, NY

Katz, G.: 2005, Attitudes toward degrees, in Maier, Bary, and Huitink (eds.), *Proceedings of SuB 9*, Radboud Universiteit Nijmegen, Nijmegen

Kennedy, C.: 2001, Polar opposition and the ontology of 'degrees', *Linguistics and Philosophy* pp 33–70

Kennedy, C.: 2005, *Vagueness and Grammar: The semantics of relative and absolute gradable predicates*, Unpublished Manuscript. Available from the author's website.

Morzycki, M.: 2004, Adverbial modification of adjectives: Evaluatives and a little beyond, in J. Dölliing and T. Heyde-Zybatow (eds.), *Event Structures in Linguistic Form and Interpretation*, Mouton de Gruyter, Berlin, To Appear

Potts, C.: 2005, *The Logic of Conventional Implicatures*, Vol. 7 of *Oxford Studies in Theoretical Linguistics*, Oxford University Press

# POLARITY ITEMS IN *BEFORE* CLAUSES

FRANCESCA PANZERI

Department of Psychology
Università Milano-Bicocca
francesca.panzeri@unimib.it

The aim of this paper is to propose a re-formulation of the uniform definition Beaver and Condoravdi (2003) proposed to account for the meaning of *before* and *after*, such that it can account also for the polarity items licensing behavior of the two temporal connectives.

## 1. Introduction

The temporal connectives *before* and *after* appear to be converses, on the one hand, but they also display different properties. In a recent paper, Beaver and Condoravdi (hereafter, B&C) proposed a uniform account of their meaning, with the intent of explaining their differences appealing to other factors – specifically, the asymmetric nature of time branching is meant to account for the different veridical properties. In that paper, nevertheless, the fact that only *before* licenses the occurrence of Polarity Items (PIs) is left unaccounted for. My aim is to show that it is possible to connect PIs licensing as well to the structure of time branching.

## 2. The data

It is well-known that *before* and *after* diverge in their logical, veridical, and licensing properties. *After* constitutes a veridical operator, inasmuch as from the truth of (1) we are entitled to infer the truth of the *after*-clause. As for *before*, it may receive a *factual* interpretation, as in (2), where the *before*-clause is implied to be true; a *non-committal* one, as in (3), where the *before*-clause is implied to have been likely when the event described in the main clause took place; and, finally, a *counterfactual* reading, as in (4), in which the *before*-clause is implied to be false:

1) Fred came home after Wilma left.
2) Fred bought a Toyota before the price went up.
3) Fred left the country before anything happened.
4) Fred died before he saw his grandchildren.

As for licensing properties, Polarity Items are licensed in *before*-clauses; and normally lead to ungrammaticality in *after*-clauses:

5) * Fred left the party after *anyone* else did.
6) Fred left the party before *anyone* else did.

## 3. Beaver and Condoravdi (2003)

Beaver and Condoravdi defend a unified account for *before* and *after*: their plot is to propose a single lexical schema and to attribute their diverging behavior to other factors. Their first step is to introduce a coercion operator, *earliest*, that ranges over times that verify a clause, and that picks up the earliest amongst them. *Before* and *after* are defined as connectives ordering a time *t* that verifies the main clause *A* with respect to the earliest time verifying the subordinated clause *B*:

> **B&C – *A after/before B* (first shot)**
> *A before/after B* is true in $t_0$ iff there is an *A*-time *t* that precedes/follows the *earliest* time that verifies the *B* clause.

As it stands, this preliminary definition cannot account for non-veridical instances of *before*-clauses, since for the truth of *A before B*, the *A*-time must precede *the earliest B-time*. B&C's solution is to exploit the definedness requirement associated with the coercion operator *earliest*: *earliest* must pick up the left boundary of an interval of times verifying the *B*-clause. If there are *no B*-times at all in the evaluation world, alternative worlds are taken into consideration.
These alternative worlds are defined as the *historical alternatives* to the evaluation world w at a time t – *alt*(w,t) – those worlds that coincide with w up to t, and from that moment may diverge only in reasonable ways, i.e., the normal future continuations of w after t. *Earliest* is thus defined relatively to this expanded domain of worlds.

> **B&C – Historical alternatives**
> *alt*(w,t) = λw'. w' is indistinguishable from w for all times t' < t; and w' is a normal continuation of w after t.

> **B&C – *A after/before B***
> $[[A \text{ after [before] } B]]^w = 1$ iff
> $(\exists t: \langle w,t\rangle \in A)\ t > [<]\ earliest.\ \lambda t'.\ (\exists w' \in alt(w,t))\ \langle w',t'\rangle \in B$

According to B&C, then, the difference between *before* and *after*'s veridical properties derives from the asymmetry of time branching: once we establish a time t (i.e., the time in which the main clause *A* holds), what is past with respect to t is fixed – and thus the set of historical alternatives to w at t is in fact reduced to the evaluation world w itself, whereas what is future with respect to t may involve different future branches, i.e., it calls for a set of historical alternative worlds.

Somehow more formally, in evaluating *A after B*, since the historical alternatives by definition coincide with w for all times t' that precede t, and since the earliest *B*-time is to be located before the *A*-time t, the set *alt*(w,t) is in fact reduced to the evaluation world alone: alt(w,t) = {w}. Thus the definition can be simplified:

**B&C – *A after B* simplified definition:**
$[[\text{A after B}]]^{w} = 1$ iff $(\exists t: \langle w,t\rangle \in A)\ t > \textit{earliest}.\ \lambda t'.\langle \mathbf{w},t'\rangle \in B$

For the sentence to be true, there must be an *A*-time t that follows the earliest amongst the times t' that verify *B in the evaluation world w*. Thus, for the sentence to be true, the subordinated clause *B* has to be instantiated in the evaluation world.
In the assessment of a *before*-sentences, the situation is different. Since the event in the *B*-clause is future with respect to the *A*-time t, historical alternatives (i.e., future branches) of w after t are activated: *B* is to be instantiated in at least one of these branches – not necessarily in the evaluation world. For instance, the sentence in (3) is predicted to be true only if something happens in one of the future continuations of the evaluation world w after the time t in which Fred left the country – and Fred's leaving must precede the earliest time in which this is the case.

## 4. The proposal

The evaluation of a *before*-clause may require considering alternative worlds; an *after*-clause is assessed with respect to the evaluation world. I propose to connect the licensing of polarity items precisely to this difference.

With a rough simplification, *A before/after B* is true if and only if there is an *A*-time *t* that precedes/follows the earliest *B*-time *t'*. My plot is to order the *A*-time t with respect to *all* the left-boundaries of intervals verifying the *B*-clause relatively to the different branches that may be activated.

***A after/before B***
$[[A \text{ after [before] } B]]^{w} = 1$ iff $\exists t\ [\langle w,t\rangle \in A\ \&$
$\forall t'\ [(t' = \textit{earliest}.\ \lambda t''.\ (\exists w' \in \text{alt}(w,t))\ \langle w',t''\rangle \in B) \rightarrow t\ > [<]\ t']]$

With this move, the initial definition of *before* and *after* renders the temporal subordinated clause a downward entailing context, i.e., a PIs licensing environment. The asymmetric nature of time branching ensures that in the assessment of *before*-clauses there may be various future branches activated, and thus different *B*-intervals to take into account – and this allows the occurrence of Polarity Items; and that the evaluation of an *after*-clause, on the other hand, is restricted to a single evaluation world, and therefore there is no need for a universal quantification over *earliest B*-times – and thus Polarity Items are not licensed.

More precisely, when *A before B* is assessed, the event in the *B*-clause follows the event in the *A*-clause, and this amounts to saying that there might be many branches in which *B* is instantiated (thus, many earliest *B*-times). In order to evaluate *A before B*, we first take into consideration all the time-world pairs <w',t''> that verify *B*, for any world w' that belongs to the set of historical alternatives to w at t; and then we collect all the times t' that are the earliest amongst them. The sentence *A before B* is true in *w* iff there is an *A*-time *t* that precedes all the earliest times *t'*. In this reformulation of the definition, the temporal clause *B* constitutes a downward entailing context:

***A before B***
$[[A \text{ before } B]]^{w} = 1$ iff $\exists t\, [<w,t> \in A \;\&$
$\forall t'\, [(t' = \textit{earliest}.\ \lambda t''.\ (\exists w' \in \text{alt}(w,t))\ <w',t''> \in B) \rightarrow t < t']]$

The initial definition for *after*-sentences mirrors the one for *before*, with only the direction of temporal ordering reversed.

***A after B* - def. 1:**
$[[A \text{ after } B]]^{w} = 1$ iff $\exists t\, [<w,t> \in A \;\&$
$\forall t'\, [(t' = \textit{earliest}.\ \lambda t''.\ (\exists w' \in \text{alt}(w,t))\ <w',t''> \in B) \rightarrow t > t']]$

But, as B&C argued, since the *B*-times t' precede the *A*-time t, the set of historical alternatives is reduced to the evaluation world, thus the definition can be simplified:

***A after B* - def. 2:**
$[[A \text{ after } B]]^{w} = 1$ iff $\exists t\, [<w,t> \in A \;\&$
$\forall t'\, [(t' = \textit{earliest}.\ \lambda t''.\ <\mathbf{w},t''> \in B) \rightarrow t > t']]$

Taking into consideration only a single world, if the *after*-clause is in fact instantiated in the evaluation world, there is a unique earliest time *t'*. Thus, there is no need to universally quantify over all the earliest *B*-times, and thus the definition can be further simplified to:

***A after B* - def. 3:**
$[[A \text{ after } B]]^w = 1$ iff
$\exists t\, [<w,t> \in A \,\&\, t > \textit{earliest}.\lambda t''.<w,t''> \in B]$

In this last simplified definition, the *after*-clause does not constitute anymore a downward entailing context (since the initial universal quantification over earliest *B*-times is reduced to a statement about the unique earliest *B*-time, because of the reduction of *alt(w,t)* to {w} itself). Thus, Polarity Items are predicted to be ungrammatical in *after*-clauses.

### 4.1. Linebarger's counterexamples

Linebarger noticed how not all instances of Polarity Items in *after*-clauses lead to ungrammaticality:

7) He kept writing novels long after he had any reason to believe they would sell.

Sentences like (7) represent a counterexample to my claim that *after*-clauses do not constitute a downward entailing context, after all the simplifications took place. But my question is: does it exist a clear criterion to identify a class of *after*-clauses that license PIs? Notice that the presence of an adequate measure phrase (such as *long*) does not constitute neither a necessary nor a sufficient condition:

8) Some say the cuts were made after there was any real use for them.
9) * He kept writing novels long after he retired to *any* Caribbean island.

Let me also highlight how the more natural Italian translation of (7) would mark the subordinated clause with subjunctive mood (the mood selected by *before*) – even if in normal *after* clauses the indicative is the only viable option:

10) Ha continuato a scrivere racconti molto dopo che ci fosse alcuna speranza.
Lit.: Has continued to write novels long after that cl. $was_{SUBJ}$ any hope.

And subjunctive mood marking is related to the activation of alternative worlds. Thus, my answer is that, even if I do not have (yet) a clear explanation of the facts, it seems to me that these kinds of sentences require the consideration of alternative branches in which the subordinated clause gets realized – even if the subordinated clause is to be placed in the past of the main clause event.

### 4.2. Beaver and Condoravdi (2004)'s proposal

In a (2004) aggregate hand-out, Beaver and Condoravdi do have a proposal for the PIs licensing: PIs are licensed in contexts that warrant strengthening inferences, provided that presuppositions are met (cf. von Fintel's notion of Strawson-entailment). *Before* and *after* diverge because the former orders the *A*-time with respect to the whole interval in which the *B* clause took place, and this is not always the case with *after*.

My objection to this line of explanation is that not all cases in which the *A*-event is ordered with respect to the whole *B*-event license PIs. For instance, an achievement predicate in the *after*-clause does not license PIs.[1] Moreover, there is evidence that the Italian counterpart of *after* (*dopo che*) always orders the main clause event with respect to the whole, completed, *B*-event – nevertheless it does not license Polarity Items. More generally, I think that it is a more efficient and natural move to resort to the same kind of explanation (i.e., the asymmetric nature of time branching) to account for both veridical and licensing properties.

**Acknowledgements**

This work was presented at Sinn und Bedeutung X (Berlin, 13-15 October 2005). I would like to thank the audience for their helpful comments. Special thanks go to Gennaro Chierchia, Carlo Cecchetto, Ivano Caponigro and Carlo Geraci, for their insightful suggestions. All errors are mine.

**Bibliography**

Anscombe, G.E.M.: 1964, Before and after, in *The Philosophical Review 73*: 3-24.

Beaver, D. & Condoravdi, C.: 2003, A Uniform Analysis of *Before* and *After*, in R. Young and Y. Zhou (eds.), *Proceedings of SALT XIII*, CLC Publications, Cornell.

Heinämäki, O.: 1974, *Semantics of English Temporal Connectives*, Ph.D. thesis, University of Indiana at Bloomington.

Landman, F.: 1991, *Structures for semantics*, Kluwer, Dordrecht.

Ogihara, T.: 1995, Non-factual *Before* and Adverbs of Quantification, in M. Simon and T. Galloway (eds.), *Proceedings of SALT V*, CLC Publications, Cornell.

[1] Notice that when achievement predicates appear in *before*-clauses, *any* is indeed grammatical – thus strengthening inferences ought to go true even if the event denoted by the predicate is punctual.

# *ALMOST*: A TEST?

DORIS PENKA

Seminar für Sprachwissenschaft
Universität Tübingen
d.penka@uni-tuebingen.de

Modifiability by *almost* has been used as a test for the quantificational force of a DP without stating the meaning of *almost* explicitly. The aim of this paper is to give a semantics for *almost* applying across categories and to evaluate the validity of the *almost* test as a diagnosis for universal quantifiers. It is argued that *almost* is similar to other cross-categorial modifiers such as *at least* or *exactly* in referring to alternatives ordered on a scale. I propose that *almost* evaluates alternatives in which the modified expression is replaced by a value close by on the corresponding Horn-scale. It is shown that a semantics for *almost* that refers to scalar alternatives derives the correct truth conditions for *almost* and explains restrictions on its distribution. At the same time, taking the semantics of *almost* seriously invalidates the *almost* test as diagnosis for the nature of quantifiers.

## 1. Introduction

Modifiability by *almost* has been used in the literature as a test for the quantificational force of a DP. At the heart of this test lies the observation hat universal quantifiers can be modified by *almost*, whereas existentials cannot. The following examples illustrate this contrast.

(1) a. Almost every student passed the exam.
b. *Almost a / some student passed the exam.

Consequently, so the argument goes, if some DP whose quantificational status is unclear can be modified by *almost*, it must have universal force. So the *almost* test has been used as an argument in the discussion of elements for which it is notoriously unclear whether they should be analysed as universals or existentials. Carlson 1981 used modifiability by *almost* to distinguish between NPI *any* and Free Choice *any* and argued that, since Free Choice *any*, but not NPI *any* can be modified by *almost*, the former is a universal quantifier, whereas the later is an existential.

(2) a. Almost any student can solve this problem set. Free Choice
b. *I didn't see almost any student. NPI

Subsequently, the *almost* test has also been used to help decide the nature of so called n-words in Negative Concord languages. Zanuttini 1991 used the fact that n-words can be modified by *almost* to argue that n-words are universal quantifiers scoping over negation, rather than existentials in the scope of negation.

(3) Non ha detto quasi niente / *alcunche. (Italian)
not have said almost n-thing / anything
'He said almost nothing.'

However, as long as the meaning of *almost* is not explicitly stated and selectional restrictions derived from it, it remains unclear what *almost* is really sensitive to and whether the arguments based an modifiability by *almost* are valid.

The aim of this paper is to give a semantics for *almost* applying across categories and to evaluate the validity of *almost* as a diagnosis for universal quantifiers under this semantics.

## 2. The meaning of *almost*

As I argued in Penka 2005, existing analysis of *almost* by Sadock 1981 and Morzycki 2001 are insufficient. They both assume that *almost* applying to a proposition p is true if p is true in a world which is not very different from the actual world. But these accounts based on intensional similarity either give wrong truth conditions for VP-modifying *almost* or cannot derive the correct selectional restrictions.

I propose that the semantics of *almost* is analogous to that of similar expressions such as *at least, at most* or *more than*. Like *almost*, these expressions can modify elements of different syntactic categories, such as adjectives, VPs and DPs:

(4) a. John was almost / at least satisfied.
b. The alpinist almost / at least reached the base camp.
c. Almost / at least half of the candidates passed the exam.

McNally 1998 and Krifka 1999 argue that expressions such as *at least, at most* or *more than* have a cross-categorial semantics similar to the semantics Rooth 1985 gives for *only*, but crucially involve alternatives ranked on a scale. Krifka assumes that these alternatives are either introduced by focus, marked by accent, or come about from expressions that are part of a Horn scale, i.e. a scale ordered by the entailment relation such that an element of the scale entails all the elements ranked lower. To ensure that the relevant alternatives are available at the level where they are evaluated, he further assumes that the scalar ordering is projected along with the focus alternatives, so that the ranking of the alternatives having the type of the focus value carries over to the alternatives at the propositional level.

For the implementation of scalar alternatives, I follow Schwarz 2005 who assumes that operators evaluating scalar alternatives have a restrictor variable ranging over scales of propositions. In the case of *almost*, the relevant alternatives are the

ones which are close by on the ordered scale. I will use ≈ to signify the 'close by'-relation and as the corresponding restrictor variable.

Here are the truth conditions I propose for *almost*:

(5) $[\![\textbf{almost}_{\approx}]\!] = \lambda w.\lambda p_{<s,t>}.\ \exists q\ [\ q \approx p\ \&\ q(w)]\ \&\ \neg p(w)$

So *almost* applied to a proposition p is true iff p itself is false in the actual world but there is an alternative proposition that is close by to p and true. There is some debate whether the requirement that p be false in the actual world is an entailment or an implicature (as argued for a.o. by Sadock 1981). I do not want to go into this discussion and will simply follow Rapp and von Stechow 1999 in assuming that it is indeed part of the truth conditions.

Note that it is only required that the alternatives under consideration be close to p, but not that they are ranked lower than p. That only alternatives ranked lower can be true is ensured by the second conjunct in (5), which requires that p be false. Since p is logically entailed by alternatives ranked higher on a Horn scale, only alternatives ranked lower can be true.

To see how this semantics works, consider the sentence in (6a), in which the scale is given by the sequence of natural numbers. Let us assume that the values that count as 'close by' are the ones within a deviation of 10% of the original value. The restrictor variable ≈ then denotes the set of propositions in (6b). Applying the meaning of *almost* stated in (5) derives the truth conditions (6c), which in effect say that the number of people who died of the disease is somewhere between 90 and 99.

(6) a. Almost 100 people died of the disease.
b. $\{p \mid p =$ that n people died of the disease, $90 \leq n \leq 110\}$
c. n people died of the disease, $90 \leq n \leq 110$ & ¬(100 people died of the disease)

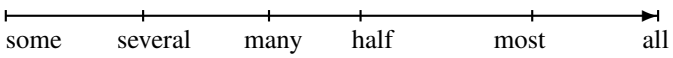
## 3. Implications for *almost* as a test

With this semantics at hand let us now see whether *almost* can indeed be used as a test for the force of a quantifier.

### 3.1. *almost* and quantifiers

As argued for by Horn 1972, quantifiers form a scale ordered by entailment:

(7)




Considering this quantifier scale we can explain why certain quantifiers cannot be modified by *almost*. We observe that vague quantifiers such as *several, many* and *most* are incompatible with *almost*, while *half* and *all* are fine:

(8) a. *Almost several / many / most students passed the exam.
b. Almost half / all of the students passed the exam.

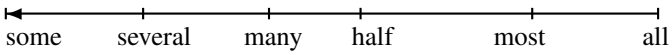
As argued by Hitzeman 1992, vague quantifiers do not correspond to precise values on the scale. Consequently it is not clear what part of the scale counts as 'close by', and so the semantics of *almost* is not compatible with vague quantifiers. In contrast, *half* and *all* have a precise location on the scale and are therefore fine with *almost*.

Furthermore, recall that existentials as in example (1b) cannot be modified by *almost*. This can be attributed to the fact that existentials form the bottom of the quantifier scale. There is thus no lower value which can be part of a true alternative as required by the semantics of *almost*.

### 3.2. n-words modified by *almost*

But the fact that existentials are at the bottom of the quantifier scale does not mean that they can never be modified by *almost*, as the *almost* test presumes. Under negation, the implication relations are reversed, leading to reversal of the direction of the corresponding Horn scale.

(9) Quantifier scale in negative contexts




Under negation, existentials are at the top of the scale. There are thus values lower on the scale which can be part of an alternative proposition that is true. Thus *almost* is not prevented from modifying existentials as long as they are in the scope of negation and *almost* operates on the negated proposition.

Since n-words in Negative Concord languages generally have an interpretation equivalent to existentials in the scope of negation, the fact that they can be modified by *almost* (cf. 3) does not imply that they are not existential quantifiers. To illustrate this consider the Italian example (10) (from Zanuttini 1991) under the proposed semantics of *almost* in combination with the assumption that *nessuno* is an existential quantifier.

(10) Non ha telefonato quasi nessuno. (Italian)
not has called almost n-person
'Almost nobody called.'

(11) {that it is not the case that a few people called, that it is not the case that a couple of people called, that it is not the case that several people called}

(12) $\exists p$ [ $p \approx$ (that it is not the case that some people called) & p] & $\neg$(that it is not the case that some people called)

In this case the restrictor variable $\approx$ denotes the set of propositions explicated in (11). Assuming that *almost* operates on the whole negated proposition we get the

truth conditions in (12) which are satisfied if somebody called, but not more than a small number of people called. This covers the meaning of (10) correctly.

Thus modifiability by *almost* does not help to decide the nature of n-words.

### 3.3. Imcompatibility of *almost* and NPIs

This leaves the question why NPIs, which are assumed to be existentials occurring in negative contexts, are not compatible with *almost* (cf. 2b and 3). I propose that this incompatibility can be reduced to intervention effects, which are known since Linebarger 1980 to arise in the licensing of NPIs.

Beck ta argues that intervention effects (in wh-questions etc.) are due to focus interpretation, or more generally the evaluation of alternative sets. An intervention effect occurs whenever an alternative evaluating operator interferes in the evaluation of another operator involving alternatives. She states this as the General Minimality Effect, which claims that the evaluation of alternatives introduced by an XP cannot skip over an intervening $\sim$ operator (i.e. the operator evaluating focus alternatives). This excludes constellations of the form (13):

(13) *[ $\mathrm{Op}_1 \ldots [\ \sim \mathrm{C}\ [\ldots \mathrm{XP}_1 \ldots ]]]$

Beck proposes that intervention effects with NPIs are also caused by the General Minimality Effect since NPI licensing is also assumed to involve alternatives (see Krifka 1995).

Under this analysis of intervention effects in NPI licensing, *almost* is predicted to be an intervener, since its semantics involves the evaluation of an alternative set. The combination of *almost* and NPIs leads to a constellation as (13), which is excluded by the General Minimality Effect.

## 4. Conclusion

In this paper I proposed a cross-categorial semantics for *almost* that is analogous to that of other similar operators such as *only*, and in particular *at least* and *more than*. According to this semantics, *almost* refers to alternatives on a Horn scale and signifies that some alternative close by on the corresponding scale is true. I showed that this semantics derives the correct truth conditions and derives the selectional restrictions observed for *almost*. Under this semantics (un)modifiability of a DP by *almost* does not tell much about the quantificational nature of the DP. In particular, the *almost* test is not a valid diagnosis for universal quantifiers.

**Acknowledgements**

For discussion and helpful comments I want to thank Arnim von Stechow, Irene Heim, Sigrid Beck, Sveta Krasikova, Ventsislav Zhechev, Hedde Zeijlstra, Torgrim Solstad, Rick Nouwen and the audience of Sinn und Bedeutung 10 in Berlin.

## Bibliography

Beck, S.: t.a., Intervention effects follow from focus interpretation, to appear in *Natural Language Semantics*

Carlson, G.: 1981, Distribution of free-choice 'any', in *Chicago Linguistics Society 17*, pp 8–23

Hitzeman, J.: 1992, The selectional properties and entailments of 'almost', in *Chicago Linguistics Society 28*, pp 225–238

Horn, L. R.: 1972, *On the Semantic Properties of Logical Operators in English*, Ph.D. thesis, UCLA

Krifka, M.: 1995, The semantics and pragmatics of polartity items, *Linguistic Analysis* 25, 209–258

Krifka, M.: 1999, At least some determiners aren't determiners, in K. Turner (ed.), *The Semantics/Pragmatics Interface from different points of view*, pp 257–291, Elseviewer Science, Oxford

Linebarger, M. C.: 1980, *The Grammar of Negative Polarity*, Ph.D. thesis, MIT

McNally, L.: 1998, Existential Sentences without Existential Quantification, *Linguistics and Philosophy* 21, 353–392

Morzycki, M.: 2001, Almost and its kin, across categories, in *Proceedings of SALT XI*, NYU

Penka, D.: 2005, *Almost there: The meaning of 'almost'*, paper presented at *Sinn und Bedeutung* 10, Berlin

Rapp, I. and von Stechow, A.: 1999, Fast "almost" and the visibility parameter for functional adverbs, *Journal of Semantics* 16, 149–204

Rooth, M.: 1985, *Association with focus*, Ph.D. thesis, University of Massachusetts at Amherst

Sadock, J.: 1981, Almost, in P. Cole (ed.), *Radical Pragmatics*, pp 257 – 271, Academic Press, NY

Schwarz, B.: 2005, Scalar additive particles in negative contexts, *Natural Language Semantics* 13, 125–168

Zanuttini, R.: 1991, *Syntactic Properties of Sentential Negation: A Comparative Study of Romance Languages*, Ph.D. thesis, University of Pennsylvania

# SEMANTICS OF POSSESSIVE DETERMINERS

STANLEY PETERS
DAG WESTERSTÅHL

Linguistics Department, Stanford University
Department of Philosophy, Göteborg University
peters csli.stanford.edu, dag.westerstahl phil.gu.se

**Abstract**

We give a uniform account of a wide range of possessive determiners, including simple (John's), quantified (few doctors'), and partitive (each of most students'), focusing on certain (frequently neglected) features of their semantics. One is the *mode of quantification* over the 'possessed' objects: often universal, but other modes are allowed too. Another is what Barker 1995 calls *narrowing*: we agree that it belongs to the semantics of possessives but note that it appears to lead to certain methodological problems. A third is the *role of definiteness* for possessives: we compare our account to the 'definiteness account' common in the literature, and in particular discuss the definiteness of partitives. Fourth, we study the *monotonicity behavior* of possessives.

## 1. Background

A (generalized) *quantifier of type* $\langle 1,1\rangle$ (of type $\langle 1\rangle$) is a mapping $Q$ that with each universe $M$ associates a binary (unary) relation $Q_M$ between subsets of $M$. Determiners typically denote type $\langle 1,1\rangle$ quantifiers, including the possessive determiners studied here, and noun phrases denote type $\langle 1\rangle$ quantifiers. Determiner denotations characteristically have the following properties:

(CONSERV) $Q_M(A,B) \Leftrightarrow Q_M(A, A\cap B)$

(EXT) for $A, B \subseteq M \subseteq M'$, $Q_M(A,B) \Leftrightarrow Q_{M'}(A,B)$

EXT applies to quantifiers of any type; in particular, many NP denotations are EXT, such as those of proper names (if John = $j$, $(I_j)_M(B) \Leftrightarrow j \in B$), bare plurals (if $C$ is the set of firemen, $(C^{\mathrm{pl}})_M(B) \Leftrightarrow \emptyset \neq C \subseteq B$), and type $\langle 1\rangle$ quantifiers $Q^A$ got by *freezing* the restriction argument of an EXT type $\langle 1,1\rangle$ quantifier $Q$ as a set $A$:

(1) $(Q^A)_M(B) \Leftrightarrow Q_{M\cup A}(A,B)$

We focus on prenominal (also known as Saxon) genitives, which can be construed as determiners. *Basic possessives* like

(2) John's, no doctors', at least five teachers', most children's

and taken to be formed by a rule

(poss) Det $\longrightarrow$ NP 's

subject to certain (light) restrictions on the NP.[1] *Complex possessives*, as in

(3) few of John's, all but five of Mary's, each of most students'

are taken to be formed by

(plex) Det $\longrightarrow$ Det of Det

where the second Det is a basic possessive.[2] (plex) applies to other Dets too (see below), though under heavy restrictions on the Dets. A main task is to provide correct and uniform truth conditions for sentences with basic and complex possessives.

## 2. Universal Readings and Others

(4) John's bikes were stolen.

usually means that each of John's bikes was stolen: a *universal* reading, i.e., with universal quantification over the 'possessed' objects. But other modes of quantification are used as well; in

(5) At most two cars' tires were slashed.

the mode is existential: at most two cars are such that *some* of their tires were slashed. (With 7 cars, the universal reading would, unreasonably, allow up to 23 slashed tires.) This indicates that the mode is given by an implicit quantifier parameter ($Q_2$) in basic possessives, whereas $Q_2$ is explicitly specified in complex possessives:

(6) Several of John's CDs were stolen.

(7) Three of each country's athletes carried a banner.

## 3. Narrowing

Consider

(8) a. Most people's grandchildren hate them.

b. Most people's grandchildren love them.

[1] We also account for possessives where a numeric expression is inserted, as in "several of John's ten", but omit them from this discussion for simplicity.

[2] An alternative is to use a rule

(part) NP $\longrightarrow$ Det of NP

Keenan and Stavi 1986 argue at length that (plex) is preferable. In our treatment, both rules have their advantages and drawbacks, but not much hinges on which one we choose.

Presumably, (8a) is false and (8b) true. Most people in the world (being too young) don't have grandchildren, but this fact is clearly irrelevant to the truth value of (8a,b), since in both cases quantification is *narrowed* to people with grandchildren. Otherwise (8a) would be trivially true, on the universal reading. Some prefer universal readings with existential import (e.g. with $\mathit{all}_{\text{ei}}(A,B) \Leftrightarrow \emptyset \neq A \subseteq B$), but, without narrowing, that has an equally undesirable effect: each of (8a,b) would then imply that most people have grandchildren. The narrowing effect was observed in Barker 1995. Although there are a few cases where narrowing seems not to be in force, and some where it doesn't affect truth conditions (see below), a vast number of sentences with possessives simply get the wrong truth conditions without narrowing.

## 4. The Possessor Relation

Semanticists agree that the choice of possessor relation is *free* in the following sense:

(Free) For any possessive NP, however predictable and semantically describable its *usual* possessor relation is, circumstances can always be found where that same possessive NP is used with *another* possessor relation, not derivable from grammatical or lexical information, but provided only by the context.

We conclude that a general treatment of possessive determiners should leave a free parameter $R$ for this relation. Further mechanisms can then describe how $R$ is (usually) fixed when it comes from, say, a relational noun like "sister".

Let $R_a = \{b : R(a,b)\}$ (the set of things $R$'d by $a$), and $\mathit{dom}_A(R) = \{a : \exists b \in A \text{ s. t. } R(a,b)\}$ (the set of objects that $R$ things in $A$).

## 5. The Meaning of Possessives

In view of the rule (poss), a semantic operator *Poss* (taken to interpret the possessive 's) should ideally take a type $\langle 1 \rangle$ quantifier $Q$ as argument, in addition to the already mentioned parameters $Q_2$ and $R$. But this makes it difficult to enforce narrowing when the possessive NP is quantified, since in general the set $C$ cannot be recovered from $(Q_1)^C$. We therefore take both $Q_1$ and $C$ as arguments, and define (for CONSERV and EXT $Q_1, Q_2$; the universe $M$ can therefore be suppressed):

(9) $\mathit{Poss}(Q_1, C, Q_2, R)(A,B) \Leftrightarrow Q_1(C \cap \mathit{dom}_A(R), \{a : Q_2(A \cap R_a, B)\})$

So quantification with $Q_1$ is narrowed to $\mathit{dom}_A(R)$ (in both arguments, by CONSERV). For certain unquantified possessive NPs we can use facts like $I_j = (\mathit{all}_{\text{ei}})^{\{j\}}$, $C^{\text{pl}} = (\mathit{all}_{\text{ei}})^C$, $I_j \vee I_m = \mathit{some}^{\{j,m\}}$; this seems somewhat *ad hoc*, but apparently suffices for all the NPs actually allowed in (poss).

The semantic rule corresponding to (plex), on the other hand, is straightforwardly compositional and merely sets the $Q_2$ parameter to the interpretation of the first Det.

This is seen to give the desired truth conditions. Forgoing narrowing would mean using $\mathit{Poss}^{\text{w}}$ instead, defined (for CONSERV and EXT $Q_1$, and EXT $Q$) by

(10) $\mathit{Poss}^{\mathrm{w}}(Q, Q_2, R)(A, B) \Leftrightarrow Q(\mathit{dom}_A(R) \cap \{a : Q_2(A \cap R_a, B)\})$

When $Q_1$ is symmetric, one sees that $\mathit{Poss}^{\mathrm{w}}((Q_1)^C, Q_2, R) = \mathit{Poss}(Q_1, C, Q_2, R)$, but in most other cases, using $\mathit{Poss}^{\mathrm{w}}$ gives the wrong result.

The chosen syntactic and semantic rules also account successfully for iterated possessive constructions, as in

(11) Mary's sisters' friends' children were there.

(12) One of John's ex-wives' previous husbands were millionaires.

(13) One of John's ex-wives' previous husbands was a millionaire.

(14) Both of many of my friends' parents work.

as well as the non-acceptability of

(15) #Many of some of John's books are stained.

## 6. Possessives and Definiteness

In the literature on possessives one often finds statements that possessive are definite (e. g. Lyons 1986, p 124, Abbott 2004, p 123). However, as soon as one goes beyond simple possessives like "John's", this is just not the case. More precisely, using the notion of definiteness from Barwise and Cooper 1981, one can show that

(16) if $Q_1$ is definite, so is $\mathit{Poss}(Q_1, C, \mathit{every}, R)$,

but when $Q_1$ is not itself definite, or when readings other than the universal one is used, the possessive is generally not definite.

A different claim, which seems quite common but is rarely spelled out in detail, is that possessives somehow 'contain' a definite. Such accounts, which we will call *definiteness accounts*, appear to use an analysis along the following lines:

(17) a. At least two of most students' books are stained.

b. For most students $x$, at least two of the books of $x$ are stained.

c. For $Q$ $x$, $Q_2$ of the $A$'s $R$'d by $x$ are $B$.

d. $Q(\{a : Q_2 \mathit{\ of\ the}(A \cap R_a, B)\})$

The locution "$Q_2$ *of the*" in the last line is interpreted out by a semantic rule for structures generated by (plex) when the final Det is definite, and one then sees that

(18) $Q(\{a : Q_2 \mathit{\ of\ the}(A \cap R_a, B)\}) \Leftrightarrow \mathit{Poss}^{\mathrm{w}}(Q, Q_2, R)(A, B)$

Definiteness accounts usually (a) do not implement narrowing, and (b) prefer an analysis of "the" using (depending on the syntactic number) either $\mathit{the}_{\mathrm{sg}}(A, B) \Leftrightarrow |A| = 1 \ \&\ A \subseteq B$ or $\mathit{the}_{\mathrm{pl}}(A, B) \Leftrightarrow |A| > 1 \ \&\ A \subseteq B$. We already commented on (a); for another example, note that using $\mathit{Poss}^{\mathrm{w}}$ for

(19) Firemen's wives worry about their husbands.

gives the undesirable consequence that all firemen are married. As to (b), we note that this too gives wrong results in many cases, and that $all_{\text{ei}}$ should be used instead. For example, even if narrowing is enforced in (19), using $the_{\text{pl}}$ would produce the entailment that firemen are bigamists![3]

Summing up, *provided* the definiteness account is amended to (a) somehow take care of narrowing, (b) use $all_{\text{ei}}$ instead of $the_{\text{sg}}$ or $the_{\text{pl}}$, and (c) allow that the implicit parameter $Q_2$ is always present, it seems to be extensionally equivalent to the account we offer here. The definiteness account uncovers at least a trace of the definite article in possessives (in the condition $A \cap R_a \neq \emptyset$, present in *Poss* as well as *Poss*$^{\text{w}}$). However, this looks like an indication of *existence* rather than of *definiteness*.

Let us come back to the restrictions on the rule (plex), which, we believe, are roughly as follows (where "partitive" means 'of the form [Det of Det]'):

(plex-restr) (i) The left Det must *not* be: basic possessive, or definite, or partitive. (ii) The right Det *must* be either basic possessive or definite; it *cannot* be partitive.

For some corroboration, consider

(20) a. few of the boys
b. each of the three girls
c. two of every student's books
d. *Mary's of the three boys
e. *the of the three boys
f. *the two of the three boys
g. *two of Mary's of the three boys
h. *two of three of Mary's girls

The standard view, however, is that only (plural) definites are allowed after [Det of]. Holding onto that view, while acknowledging that possessives are usually not themselves definite, requires some version of the definiteness account of possessives; we saw in (17) how this account analyzes "$Q_2$ of $Q_1$ $C$" on a form containing instead "$Q_2$ of the $C$", where a definite is indeed following [Det of]. But another, and perhaps simpler, idea is instead to revise the standard view along the lines of (plex-restr) above, which allows *both* definites and (basic) possessives after [Det of].[4]

[3] Other examples show that it also doesn't work to use $the_{\text{sg}}$ for many 'singular' cases.

[4] The semantic rule corresponding to (plex) indicated above does not work when the second Det is not a possessive, so a separate rule for this case is needed. The use of these two rules can be seen in

(i) Two of the ten boys' books are missing.

It is structurally ambiguous whether "two" quantifies over boys or books. But in the latter case, there are still two possibilities. If the rule for possessives is used, each boy is missing two books, so up to twenty books are missing in all. But if the rule for definites is used, only two books are missing, among books owned by any of the boys. Each of these three readings seems entirely plausible.

## 7. Possessives and Monotonicity

A type $\langle 1,1\rangle$ quantifier $Q$ is MON↑ (MON↓) if $Q_M(A,B)$ and $A \subseteq A' \subseteq M$ ($A' \subseteq A$) implies $Q_M(A')$. It is ↑MON or *persistent* (↓MON or *anti-persistent*) if the corresponding holds for the left argument. The following left properties are also useful: $Q$ is $\uparrow_{\text{SE}}$MON ($\downarrow_{\text{NW}}$MON) if $Q_M(A,B)$ & $A \subseteq A' \subseteq M$ ($A' \subseteq A$) & $A - B = A' - B$ implies $Q_M(A',B)$, and it is $\downarrow_{\text{NE}}$MON ($\uparrow_{\text{SW}}$MON) if $Q_M(A,B)$ & $A' \subseteq A$ ($A \subseteq A' \subseteq M$) & $A \cap B = A' \cap B$ implies $Q_M(A',B)$. $Q$ is *smooth*, if it is $\downarrow_{\text{NE}}$MON and $\uparrow_{\text{SE}}$MON, and *co-smooth*, if it is $\downarrow_{\text{NW}}$MON and $\uparrow_{\text{SW}}$MON. One can show that, under CONSERV, smoothness implies MON↑. In fact, almost all MON↑ determiner denotations are smooth, so (co-)smoothness seems to be a highly significant property for natural language quantifiers.

Possessive determiners provide a rich source of quantifiers with various monotonicity properties, usable e. g. to test hypotheses about how monotonicity relates to other linguistic phenomena, in particular to the distribution of polarity items.

The monotonicity properties of *Poss*$(Q_1, C, Q_2, R)$ are determined by those of $Q_1$ and $Q_2$ in interesting ways. For right monotonicity we have:

(21) If $Q_1$ and $Q_2$ are right monotone in the same (opposite) direction, it holds that *Poss*$(Q_1, C, Q_2, R)$ is MON↑ (MON↓).

Left monotonicity yields too many cases to describe here; we mention just one:

(22) Let $Q_2$ be ↓MON↑ and co-symmetric [i.e. $Q_2\neg$, defined by $Q_2\neg(A,B) \Leftrightarrow Q_2(A, A-B)$, is symmetric], and $Q_1$ be smooth and positive [i.e. $Q(A,B) \Rightarrow A\cap B \neq \emptyset$]. Then *Poss*$(Q_1, C, Q_2, R)$ is weakly ↓MON↑ and weakly smooth.[5]

Example: *most professors'* (universal reading).

**Acknowledgements**

We thank Barbara Partee for helpful comments, in particular in relation to the definiteness account of possessives.

**Bibliography**

Abbott, B.: 2004, Definiteness and indefiniteness, in L. R. Horn and G. Ward (eds.), *Handbook of Pragmatics*, Blackwell, Oxford, pp 122–149.

Barker, C.: 1995, *Possessive Descriptions*, CSLI Publications, Stanford.

Barwise, J. and Cooper, R.: 1981, Generalized quantifiers and natural language, *Linguistics and Philosophy* 4, pp 159–219.

Keenan, E. and Stavi, J.: 1986, A semantic characterization of natural language determiners, *Linguistics and Philosophy* 9, pp 253–326.

Lyons, C.: 1986, The syntax of English genitive constructions, *Journal of Linguistics* 22, pp 123–143.

[5] *Weak* versions of left downward monotonicity properties add the condition $C \cap dom_A(R) = C \cap dom_{A'}(R)$ in the antecedent; cf. the existence requirement for possessives mentioned above.

# DETERMINERS IN ASPECTUAL COMPOSITION

CHRISTOPHER PIÑÓN

Research Institute for Linguistics
Hungarian Academy of Sciences
pinon@nytud.hu

A difficulty for leading theories of aspectual composition (Krifka 1992; Verkuyl 1993) is that they make incorrect predictions when verbs of gradual change combine with certain kinds of noun phrases. Consider, for instance, the sentences in (1) and (2), which all describe accomplishments, as is confirmed by the acceptability of the time-span *in*-adverbials.

(1) a. Rebecca ate an apple (in five minutes).
b. Rebecca ate the apple (in five minutes).
c. Rebecca ate (exactly) three apples (in fifteen minutes).

(2) a. Rebecca ate at least three apples (in fifteen minutes).
b. Rebecca ate more than three apples (in fifteen minutes).
c. Rebecca ate some apples (in fifteen minutes).
d. Rebecca ate a number of apples (in fifteen minutes).
e. Rebecca ate at most three apples (in fifteen minutes).
f. Rebecca ate less than three apples (in fifteen minutes).

In Krifka's approach, the verb phrases and sentences in (1)—or more precisely, their corresponding event predicates, though I will often speak loosely—are predicted to be *quantized*. The definition of quantized reference for one-place predicates $P$ is given in (3), where $a$, $b$ are unsorted variables for individuals and $\sqsubset$ stands for the proper part relation.

(3) $\mathsf{QUA}(P) \stackrel{\text{def}}{=} \forall a \forall b[(P(a) \wedge P(b)) \rightarrow \neg(a \sqsubset b)]$ ($P$ is quantized)

The verb phrases and sentences in (1) are quantized because the noun phrases *an apple*, *the apple*, and *(exactly) three apples*, which are treated as nominal predicates in Krifka's framework, are all quantized, and this leads to the result (given the thematic properties of verbs of gradual change) that the corresponding verb phrases and sentences are also quantized. If *in*-phrases require a quantized verb phrase to combine with, then the acceptability of the sentences in (1) is accounted for.

In contrast, the acceptability of the sentences in (2) is problematic. The noun phrases *at least three apples*, *more than three apples*, *some apples*, and *a number of apples* in (2a)–(2d) are apparently[1] *cumulative*, hence not quantized, which pre-

[1] I write 'apparently' because on an intuitively straightforward analysis of these noun phrases as nominal predicates they are cumulative. Of course, there may be other (arguably, less intuitively straightforward)

dictably results in event predicates that are cumulative and not quantized. The definition of cumulative reference for one-place predicates $P$ is shown in (4), where $\oplus$ designates the sum operation.

(4) $\textsc{cum}(P) \stackrel{\text{def}}{=} \exists a \exists b[P(a) \wedge P(b) \wedge \neg(a = b)] \wedge \forall a \forall b[P(a) \wedge P(b) \rightarrow P(a \oplus b)]$
($P$ is cumulative)

Furthermore, *at most three apples* and *less than three apples* in (2e) and (2f) are apparently neither cumulative or quantized,[2] which means that the corresponding event predicates are also neither cumulative nor quantized. If *in*-phrases select for a quantized event predicate, then they should not be compatible with the sentences in (2), and yet they are. Observe that the same difficulty arises with noun phrases such as *many apples* (cumulative, hence not quantized) and *few apples* (neither cumulative nor quantized).

This difficulty becomes even more acute in light of the following contrast:

(5) a. Rebecca ate at least one apple in ten minutes.
b. #Rebecca ate apples in ten minutes.

Assuming that singular apples are the minimal elements in the extension of *apples*, then *apples* and *at least one apple* have the same extensions, and both are cumulative. In fact, they even have the same intensions, because it is impossible for a number of apples to satisfy one of these predicates without also satisfying the other.[3] Yet if correct, then this suggests that Krifka's strategy of looking solely at the reference properties of the nominal predicates in question is insufficiently general and really only succeeds with 'well-behaved' predicates such as *an apple*, *the apple*, and *(exactly) three apples*.[4]

Although Krifka's theory is formulated in an event semantic framework, this difficulty arises in Verkuyl's approach as well, which dispenses with events. Since the latter also relies on extensions, it is not feasible to distinguish *apples* from *at least one apple* in an aspectually relevant way. Verkuyl's claim that *apples* is $[-\text{SQA}]$ and *at least one apple* is $[+\text{SQA}]$ (SQA = 'specified quantity of $A$, where $A$ is the denotation of the head noun') is effectively to postulate a difference without a difference, because at the level of set theory there is no difference—in both cases, the set of apples in question has an unspecified positive cardinality. (Verkuyl 1993, sects. 4.3, 6.3) is clearly a bit troubled by this problem, but his ultimate appeal (on my reading) is to a difference in *representations*, a move that is strikingly incongruous with his

analyses, e.g., the one proposed in (Zucchi and White 2001), on which they are not cumulative.

[2] Since quantized and cumulative reference form contraries and not contradictories, a predicate may be neither quantized nor cumulative.

[3] On the assumption that the minimal elements are sums consisting of two apples, just replace *at least one apple* in (5a) with *at least two apples*.

[4] Note that if the *semantics* of *three apples* is really 'at least three apples', as is sometimes assumed (with the sense of 'exactly' derived as an implicature), then *three apples* would really be cumulative and not quantized. In this case, *three apples* and comparable noun phrases with a numeral would not be so 'well-behaved' and would pose basically the same problem as the object noun phrases in (2a)–(2d).

otherwise strict model-theoretic regimen.[5]

**A New Analysis: Alternatives for Determiners** The leading idea behind the new analysis is that the exact determiner chosen in constructing a verb phrase should not necessarily affect the 'quantized character' of the verb phrase as long as there is an alternative in the focus semantic value of the determiner that entails it and that would have yielded a quantized verb phrase. Here, 'alternative' and 'focus semantic value' are used as in alternative semantics for focus, as described in (Rooth 1992). It is this consideration of alternatives that distinguishes the present approach from the purely extensional approaches to aspectual composition advocated by Krifka and Verkuyl.

To set the stage informally, recall the sentence in (2a) with the object noun phrase *at least three apples*. Granting that this noun phrase is cumulative (hence not quantized), then it follows that the verb phrase is also cumulative. However, since the *in*-phrase is acceptable, we need to account for what distinguishes this case from the one in (5b), where the *in*-phrase is not acceptable even though the bare plural *apples* is cumulative as well. Adopting the speaker's perspective, the crucial choice is whether or not to use a determiner to express information about cardinality (or more generally, about quantity). If a determiner is used, then the (positive) choice is to express information about cardinality; if no determiner is used, then the (negative) choice is not to. But even if the choice is positive, the speaker may or may not know the precise cardinality of the set in question, and even if she knows it, in a given context it may not be relevant for her to express it precisely. However, the vital point is that the aspectual value of the verb phrase *should not depend on* whether or not the speaker is able to or decides to express the precise cardinality as long as she chooses to express information about cardinality in the first place. Thus, in choosing *at least three apples* in (2a), the speaker opts to use a determiner to express information about cardinality, but she might well have chosen *(exactly) four apples* had she known the precise cardinality and had it been relevant for her to express it.

In what follows, I will sketch the implementation of the new analysis in an event semantic framework similar in spirit to Krifka's. The first step is to assign determiners a much more prominent role than they play in his approach. In an event semantics, determiners are of the type $\langle\langle e,t\rangle,\langle\langle e,\langle\varepsilon,t\rangle\rangle,\langle\varepsilon,t\rangle\rangle\rangle$, where $e$ is the type of ordinary individuals and $\varepsilon$ is the type of events. The determiner applies first to a predicate $P$ of ordinary individuals and then to a relation $R$ between events and ordinary individuals, yielding an event predicate, e.g.:

(6) a. a(n) $\rightsquigarrow \lambda P\lambda R\lambda e[\exists x[R(e,x)\wedge P(x)\wedge |x|=1]]$
b. (exactly) three $\rightsquigarrow \lambda P\lambda R\lambda e[\exists x[R(e,x)\wedge P(x)\wedge |x|=3]]$
c. at least three $\rightsquigarrow \lambda P\lambda R\lambda e[\exists x[R(e,x)\wedge P(x)\wedge |x|\geq 3]]$
d. at most three $\rightsquigarrow \lambda P\lambda R\lambda e[\exists x[R(e,x)\wedge P(x)\wedge |x|\leq 3]]$

The second step is to define notions of quantized and cumulative reference for de-

[5] Unfortunately, due to the lack of space, I cannot review the various proposals in (Krifka 1998; Zucchi and White 2001; Rothstein 2004; Borer 2005) for this difficulty, but I hope to do so on another occasion.

terminers, because the ones in (3) and (4) are not directly applicable to determiners. The notion of quantized reference for determiners is *quantized*$^*$, a predicate of (the meanings of) determiners $\mathscr{D}$:

(7) $\mathsf{QUA}^*(\mathscr{D}) \stackrel{\text{def}}{=} \forall e \forall R \forall P \forall x [$
$\mathscr{D}(e,R,P) \wedge \mathsf{GRAD}(R) \wedge \neg \mathsf{ITER}(e,x,R) \wedge \mathsf{SUM}(R) \wedge \mathsf{CUM}(P) \rightarrow$
$\mathsf{QUA}(\lambda e'[\mathscr{D}(e',R,P)])]$ ($\mathscr{D}$ is quantized$^*$)

A determiner $\mathscr{D}$ is quantized$^*$ just in case for every $e$, $R$, $P$, and $x$ the following holds: if $\mathscr{D}$ applies to $e$, $R$, and $P$, $R$ is *gradual* (GRAD), not *iterative* (ITER) with respect to $e$ and $x$, and *summative* (SUM), and $P$ is cumulative, then the event predicate resulting from the application of $\mathscr{D}$ to $R$ and $P$ is quantized. Less formally, $\mathscr{D}$ is quantized$^*$ just in case it yields a quantized event predicate when applied to a gradual relation $R$ that is summative but not iterative and a cumulative nominal predicate $P$.[6] The notions of graduality, iterativity, and summativity are taken from (Krifka 1992, pp. 42, 40, 39) and characterize two-place relations between events and incremental patients that are summative but such that no part of the patient is affected more than once. In (6), the determiners *a(n)* and *(exactly) three* are quantized$^*$ but *at least three* is not.

The corresponding notion of cumulative reference for determiners (*cumulative*$^*$) is defined as follows:

(8) $\mathsf{CUM}^*(\mathscr{D}) \stackrel{\text{def}}{=} \exists e \exists R \exists P \exists e' [\mathscr{D}(e,R,P) \wedge \mathscr{D}(e',R,P) \wedge \neg(e = e')] \wedge$
$\forall e \forall R \forall P [(\mathscr{D}(e,R,P) \wedge \mathsf{SUM}(R) \wedge \mathsf{CUM}(P)) \rightarrow \mathsf{CUM}(\lambda e'[\mathscr{D}(e',R,P)])]$
($\mathscr{D}$ is cumulative$^*$)

A determiner $\mathscr{D}$ is cumulative$^*$ just in case there are at least two events $e$ and $e'$ in its extension and for every $e$, $R$, and $P$ the following holds: if $\mathscr{D}$ applies to $e$, $R$, and $P$, $R$ is summative, and $P$ is cumulative, then the event predicate resulting from the application of $\mathscr{D}$ to $R$ and $P$ is cumulative. In (6), *at least three* is cumulative$^*$, whereas *a(n)* and *(exactly) three* are not. Observe that if a determiner is cumulative$^*$, then it is not quantized$^*$, but that a determiner may also be neither cumulative$^*$ nor quantized$^*$ (e.g., *at most three* in (6d)).

If a determiner $\delta$ is focused (with a syntactic representation as $[_{\mathrm{D}}\ \delta]^{\mathrm{f}}$), then its focus semantic value is the (typically contextually restricted) set of alternatives of the same type as its ordinary semantic value and from which its ordinary semantic value is drawn. If $\delta$ is focused and its ordinary semantic value is represented by $\mathscr{D}$, then let's designate its focus semantic value by $\mathsf{ALT}(\mathscr{D})$, where ALT is a function that maps $\mathscr{D}$ to the set of its alternatives. In this case, a natural subset of $\mathsf{ALT}(\mathscr{D})$ is the one consisting of those alternatives $\mathscr{D}'$ that entail $\mathscr{D}$:

(9) $\mathsf{ALT}_{\subseteq}(\mathscr{D}) \stackrel{\text{def}}{=} \{\mathscr{D}' | \mathscr{D}' \in \mathsf{ALT}(\mathscr{D}) \wedge \mathscr{D}' \subseteq \mathscr{D}\}$
(the set of alternatives of $\mathscr{D}$ that entail $\mathscr{D}$)

[6] The idea is that the nominal predicates $P$ are all cumulative and it is the determiner that adds the information about cardinality. Similarly, the verbal relations $R$ are all taken to be summative (summativity is simply cumulativity for two-place relations).

For example, if *at least three* is focused, then the set of its alternatives that entail it include *(exactly) four* and *at least five* but not *(exactly) two* or *less than three* even though the latter two are included in the set of its alternatives. Note that if a determiner $\delta$ is not focused and its ordinary semantic value is represented by $\mathscr{D}$, then its focus semantic value is simply the singleton consisting of its ordinary semantic value, namely, $\{\mathscr{D}\}$.

The next step is to introduce a notion of quantized reference for determiners that is sensitive to their alternatives. A determiner $\mathscr{D}$ is *f-quantized*$^*$ just in case there is an alternative in the set of its alternatives that entail it which is quantized$^*$:

(10) $\textsf{F-QUA}^*(\mathscr{D}) \stackrel{\text{def}}{=} \exists\mathscr{D}'[\mathscr{D}' \in \textsf{ALT}_{\subseteq}(\mathscr{D}) \wedge \textsf{QUA}^*(\mathscr{D}')]$ ($\mathscr{D}$ is f-quantized$^*$)

Clearly, if $\mathscr{D}$ is quantized$^*$, then it is f-quantized$^*$, even if its syntactic correspondent is not focused (in which case $\mathscr{D}$ is the sole element of $\textsf{ALT}_{\subseteq}(\mathscr{D})$). However, the converse does not hold, because $\mathscr{D}$ may be f-quantized$^*$ without being quantized$^*$, provided that its syntactic correspondent is focused. For example, if *at least three* is focused, then it is f-quantized$^*$ but not quantized$^*$—indeed, it is cumulative$^*$ whether focused or not. In fact, all of the other 'problematic' determiners in (2) are similarly f-quantized$^*$ if focused but not quantized$^*$.

The following definition provides a corresponding notion of f-quantized reference for one-place predicates $P$:

(11) $\textsf{F-QUA}(P) \stackrel{\text{def}}{=} \exists P'[P' \in \textsf{ALT}_{\subseteq}(P) \wedge \textsf{QUA}(P')]$ ($P$ is f-quantized)

A predicate $P$ is f-quantized just in case there is an alternative in the set of its alternatives that entail it which is quantized. For example, the verb phrase *eat* $[_{\text{D}}$ *at least three*$]^{\text{f}}$ *apples*, although not quantized is f-quantized, precisely because $[_{\text{D}}$ *at least three*$]^{\text{f}}$ is f-quantized$^*$.

We can now finally state the selectional requirement of *in*-adverbials:

(12) An *in*-adverbial selects for a verb phrase that is represented by an event predicate which is f-quantized.

Let's consider how this requirement accounts for the data in (1), (2), and (5). Since the verb phrases in (1) are all quantized (due to the fact that the respective determiners are quantized$^*$), they are perforce f-quantized and no appeal to focus is necessary. In contrast, since none of the verb phrases in (2) or (5a) are quantized (due to the fact that none of the respective determiners are quantized$^*$), the only way for them to be f-quantized is for the respective determiners to be f-quantized$^*$, which means that the determiners have to be focused. Indeed, the natural way of uttering the sentences in (2) and (5a) (especially in conjunction with the *in*-adverbials) is with an intonational focus on the respective determiners.

This analysis has the nice consequence that if the noun phrase lacks a determiner (or at least an overt determiner), as in the case of bare plurals (e.g., *apples* in (5b)), then there is no determiner to focus and so only the contribution of the bare (cumulative) noun phrase can be considered, which yields a cumulative verb phrase.

A similar consequence holds for languages with an overt determiner that cannot be focused, e.g., French *des*, as in *des pommes* 'apples'. Since *des* is cumulative* (assuming that we want to interpret *des*), it would need to be focused in order to be f-quantized*, but since it cannot be focused, it cannot be f-quantized*, and thus it contributes to a verb phrase that is cumulative and not f-quantized. This correctly predicts that such verb phrases will be incompatible with *en*-adverbials (the French equivalent of *in*-adverbials), as seen in (13a) (cf. (5b)). In constrast, the determiner *quelque* 'some', although cumulative* as well, can nevertheless be focused and hence f-quantized*, which would predictably contribute to a verb phrase that is f-quantized (despite being cumulative), thereby satisfying the requirement of *en*-adverbials, as shown in (13b) (cf. (2c)).

(13) a. #Juliette a mangé des pommes en dix minutes.
Juliette has eaten ∅ apples in ten minutes
b. Juliette a mangé quelques pommes en dix minutes.
Juliette has eaten some apples in ten minutes

In sum, in its appeal to alternatives for determiners, the new analysis aims to capture the intuition that the 'problematic' determiners in (2) are merely less precise ways of expressing information about cardinality that could have been made more precise, and that the aspectual value of the verb phrase (in terms of f-quantized reference) is not necessarily affected by whether a less or more precise determiner is chosen.[7]

## Bibliography

Borer, H.: 2005, *Structuring Sense: The Normal Course of Events*, Vol. 2, Oxford University Press

Krifka, M.: 1992, Thematic relations as links between nominal reference and temporal constitution, in I. Sag and A. Szabolcsi (eds.), *Lexical Matters*, pp 29–53, CSLI Publications, Stanford

Krifka, M.: 1998, The origins of telicity, in S. Rothstein (ed.), *Events and Grammar*, pp 197–235, Kluwer Academic Publishers

Rooth, M.: 1992, A theory of focus interpretation, *Natural Language Semantics* 1, 75–116

Rothstein, S.: 2004, *Structuring Events: A Study in the Semantics of Lexical Aspect*, Blackwell Publishing

Verkuyl, H. J.: 1993, *A Theory of Aspectuality: The Interaction between Temporal and Atemporal Structure*, Cambridge University Press

Zucchi, S. and White, M.: 2001, Twigs, sequences and the temporal constitution of predicates, *Linguistics and Philosophy* 24, 223–270

[7]I wish to thank Fabienne Martin for useful discussions of this topic. This work was supported by the Hungarian Scientific Research Fund (OTKA TS 049873). The author's web page is ⟨http://pinon.sdf-eu.org⟩.

# SCOPE DISAMBIGUATION BY ELLIPSIS AND FOCUS WITHOUT SCOPE ECONOMY

MATS ROOTH

Linguistics and CIS
Cornell University
rooth@cornell.edu

This paper reconstructs the analysis in Danny Fox's *Economy and Semantic Interpretation* of ellipsis/focus scope disambiguation effects in a way which eliminates reference to scope economy, instead relying only on focus theory to constrain representations.

## 1. Introduction

In isolation, sentence (1) has a surface scope reading in which the existentially quantified subject [a boy] has scope over the quantified object [every teacher], and an inverted scope reading in which the quantified object has widest scope. Sag 1975 pointed out that verb phrase ellipsis disambiguates such sentences in the direction of surface scope. In (2), where there is verb phrase ellipsis in the second sentence, the first sentence has only surface scope.

(1) A boy admires every teacher. $(\exists > \forall)(\forall > \exists)$
(2) A boy admires every teacher. $\text{Mary}_{\text{F}}$ does too. $(\exists > \forall)$
(3) A boy admires every teacher. $\text{Mary}_{\text{F}}$ admires every teacher too. $(\exists > \forall)$

The same is observed in (3) where there is a focus on the subject *Mary*, with prosodic reduction of the verb phrase *admires every teacher*, but without ellipsis. The effect seems stronger with ellipsis, though.

Quantifier lowering constructions are a second context. In isolation, the existential quantifier contributed by *a Latvian* in the first sentence of (4) can have maximal scope, or scope under *certain*. The latter is the only reading of (5). Ellipsis and focus in the second sentence of (4) disambiguate the first sentence in the direction of surface scope. The same is seen with the stripping ellipsis in (6), though curiously, only if it is taken for granted that Paul is not a Latvian.

(4) A Latvian is certain to be at the conference. $\text{Paul}_{\text{F}}$ is, too. $(\exists > \textit{certain})$
(5) There is certain to be a Latvian at the conference. $(\textit{certain} > \exists)$
(6) A Latvian is certain to be at the conference, but not Paul.

Hirschbühler 1982 pointed out examples where ellipsis fails to disambiguate scope. (7) has a reading where *every building* takes scope in the first conjunct

over the subject *a Canadian flag*, and in the second conjunct the universal quantifier (which is part of the elided VP) takes scope over *an American flag*. The data in (8) from Fox 1999 are similar. Unlike (2), (8) has a reading where *a boy* has minimal scope.

(7) A Canadian flag is in front of every building, and an American flag is too.
(8) A boy admires every teacher. A girl does too.

A further datum was pointed out in Fox 1995. In (9a), the subject in the antecedent sentence is non-quantified, while the subject in the ellipsis sentence is quantified, reversing the situation in (2). In this case, scope remains ambiguous in the ellipsis sentence: the ellipsis sentence can have the reading (9b). The same behavior is seen in quantifier-lowering versions (10).

(9) a. Mary admires every teacher. A boy does, too.
    b. Every teacher has the property of being admired by a boy.
(10) a. Paul is certain to be at the conference, and a Latvian is, too.
     b. Paul is certain to be at the conference, as is a Latvian.
     c. Paul is certain to be at the conference, but not a Latvian.

## 2. Assumptions

Fox suggested that constraints coming from intonational focus are responsible for the paradigm. Following Tancredi 1992 and Rooth 1992a, the grammatical representation for the ellipsis sentence includes a focus feature with scope over the ellipsis, with the antecedent for the focus being some clause containing the antecedent for the ellipsis. The motivation for this is independent, having to do with non-local cases of covariant/sloppy readings for pronouns (11).

(11) $\text{John}_F$ asked Mary to help him, and $\text{Bill}_F$ asked $\text{Sue}_F$ too.
Covariant: Bill asked Sue to help Bill.

In addition to contraints coming from focus, Fox appeals to a scope economy principle, according to which quantifier raising or lowering is blocked in cases where the operation has no semantic effects, as is true for instance for proper names. The project for the present paper is to show that, if one develops the consequences of focus semantics more thoroughly than Fox does, reliance on scope economy can be dropped.

I'll assume the notation of Rooth 1992b, which represents antecedents for focus with indexing:

(12) [Franklin wrote $\text{it}_5$]$_1$ No, [$\text{Jefferson}_F$ wrote $\text{it}_5$]$\sim$1

The operator $\sim$1 marks the scope of the focus, and an antecedent. Semantically, this enforces a constraint among the proposition denoted by [$\text{Jefferson}_F$ wrote $\text{it}_5$], the

focus alternative set for that constituent (which relative to an assignment $g$ is a set of propositions of the form '$y$ wrote $g(5)$') and the proposition denoted by [Franklin wrote it$_5$]. Different formulations of this constraint have been given (Rooth 1992b, Rooth 1996, Schwarzschild 1999). Here I will assume Schwartzschild's version, which is that the antecedent entails the union of the focus alternative set.[1] Assuming that Franklin is one of the alternatives to Jefferson, this focus constraint is satisfied in (12).[2] According to the hypothesis of focus licensing of ellipsis, the compositional representation in an ellipsis version is the same, with a focus feature having scope over the ellipsis site, and the antecedent for focus having scope over the ellipsis antecedent.

Finally, we require mechanisms for ellipsis and quantifier scope which in interaction do not constrain scope too much. For instance, we do not want the ellipsis mechanism to enforce identity of first-order properties, because this would always produce maximal scope for the subject of the ellipsis property. Fox proceeds structurally, with tree transformations of quantifier raising and lowering feeding semantic interpretation and the grammar of focus. It is not completely clear to me, though, how the tree transformations interact with the grammar of ellipsis and focus. As a clean baseline theory, I will assume that VPs such as *certain to be at the conference* denote higher-order properties such as (13a). If the higher order property combines directly with a generalized quantifier (13b), narrow scope for the quantifier results. To obtain wide scope for the quantifier, the predicate is first modified with the operator (13c), which turns the higher-order property into a first-order one. The advantage of this setup is that (13c) can be assumed to be optionally present outside the VP in the antecedent and ellipsis clauses, producing any of four combinations of scopes while using a single compositional semantics for the VP.

(13) a. $\lambda\mathcal{P}\mathbf{certain}(\mathcal{P}(\mathbf{atconference}))$
b. $\lambda Q\exists y\,[\mathbf{Latvian}(y) \wedge Q(y)]$
c. Op $= \lambda\Phi\lambda x\Phi(\lambda P P(x))$

Examples like (2) will be treated similarly, using a VP denoting a higher-order property.

## 3. Analysis in Focus theory

(14a,b) gives two representations for the antecedent sentence in (4) on these assumptions, and (14c) is the representation of the ellipsis sentence. In the latter, Op is required to lower the higher-order property to a first-order property which can combine with the type $e$ subject [Paul$_F$].

[1] The reason is that I need to license the representation (14b,c).

[2] Schwarzschild does not use an indexing notation, instead stating a constraint that some antecedent must be present. He proposes that all non F-marked nodes have antecedents, so that in (12), *it$_5$*, *wrote*, and *wrote it$_5$* have antecedents.

(14) a. [[a Latvian] is [certain to be at the coference]]$_1$ *certain* $>\exists$
b. [[a Latvian] Op is [certain to be at the coference]]$_1$ $\exists >$ *certain*
c. [Paul$_F$ Op [certain to be at the coference]]$\sim$1

As required by focus licensing of ellipsis, the first sentence (either (14a) or (14b)) is the antecedent for the focus in the ellipsis sentence. This is indicated by indexing using index 1. To check whether the representations satisfy the focus constraint, we compute the focus alternative set for the argument of $\sim$. Allowing all persons as substitutes for Paul, this is (15a). The union of this set is (15b), the proposition 'someone has the property of being certain to be at the conference'. Since this is entailed by 'some Latvian has the property of being certain to be at the conference but not 'it is certain that there will be some Latvian at the conference', the discourse (14b,c) but not the discourse (14a,c) is licensed by focus theory. So as desired, an application of focus theory predicts disambiguation in the direction of surface scope.

(15) a. $\{Op(\lambda\mathcal{P}\mathbf{certain}(\mathcal{P}(\mathbf{atconference})))(y)|\mathbf{person}(y)\}$
b. $\exists y\,[\mathbf{person}(y) \wedge \mathbf{certain}(\mathbf{atconference}(y))]$

Fox's analysis proceeds as follows. Optional quantifier lowering generates a representation Ai with narrow scope for the subject in the antecedent, in addition to the surface scope representation As. Optional lowering of the subject in the ellipsis sentence produces a representation Ei, alongside the surface representation Es. This gives $2 \times 2 = 4$ combinations; a focus constraint and scope economy are checked in each of them, resulting in the table on the left in (16). Economy is satisfied in each of the surface representations, because there is no movement; it is satisfied in Ai, because quantifier lowering produces a different meaining; it is violated in Ei, because [$_{NP}$Paul] is scopeless, so lowering does not produce a new reading. Fox assumes that the effect of focus interpretation is to enforce isomorphic structural scope for the subject in the antecedent and the ellipsis clauses. This places stars in the lower left and upper right cells of the table, where structural scopes are non-isomorphic. Only the upper left cell has no violations. This is a cell with surface scope for the antecedent, so as desired the analysis predicts disambiguation in the direction of surface scope.

(16)

| Fox | Es | | Ei | |
|---|---|---|---|---|
| As | A econ | ok | A econ | ok |
| | E econ | ok | E econ | * |
| | focus | ok | focus | * |
| Ai | A econ | ok | A econ | ok |
| | E econ | ok | E econ | * |
| | focus | * | focus | ok |

| Here | Es | | Ei | |
|---|---|---|---|---|
| As | focus | ok | focus | ok |
| Ai | focus | * | focus | * |

This part of Fox's analysis is isolated from research on focus interpretation, because it does not employ a semantics for focus. Focus semantics *can* result in isomorphic scope of quantifiers being imposed, but this is not the semantics of focus.

Morover, once the gap is corrected, scope economy proves to be redundant. The focus semantic value of both Es and Ei is (15b), so that if we apply focus semantics in Fox's syntactic representations, disambiguation works out exactly as in my representation (14). So, also on Fox's syntactic assumptions, scope economy can be dispensed with.

Let us turn to Fox's observation about the symmetric case, where the subject is non-quantified in the antecedent, but quantified in the ellipsis sentence. In (17), the intuition is that scope is ambiguous in the ellipsis sentence (17c). Fox suggested that, at least in the reading with narrow scope for the subject, the focused element is *Latvian*. This tends to be supported by the fact that the discourse context makes clear that Paul is a Finn. Further, the discourse as a whole seems to topicalize an opposition between a Finn being at the conference and a Latvian being there, or generalizing, to suggest a topic 'people of what nationalities will be at the conference?'.

(17) a. Will any Finns be there?
b. Paul is certain to be at the conference.
c. A Latvian$_{\mathrm{F}}$ is, too.

(18b) is my canonical representation for the inverted scope reading of the ellipsis clause, and (18c) is my representation for the surface scope reading. (18a) is the antecedent sentence, which as before is treated as the antecedent for the focus. Concentrating on the inverted representation (18b), (19a) is the focus semantic value, where $A(Q, \mathbf{Latvian})$ expresses '$Q$ is one of a set of contextually given alternatives to $\mathbf{Latvian}$'. (19b) is the union of the focus semantic value or focus closure. On the assumption that **Finn** is one of the alternatives to **Latvian**, the focus closure is entailed by the meaning of the antecedent (18a), together with the contextual assumption that Paul is a Finn (and still will be a Finn at the time of the conference). Therefore the focus constraint is satisfied in the discourse (18a,b).

(18) a. [Paul Op is [certain to be at the coference]]$_1$
b. [[a Latvian$_{\mathrm{F}}$] is [certain to be at the coference]]$\sim$1 *certain* $>\exists$
c. [[a Latvian$_{\mathrm{F}}$] Op is [certain to be at the coference]]$\sim$1 $\exists >$ *certain*

(19) a. $\{\mathbf{certain}(\exists y\,[Q(y) \wedge \mathbf{atconference}(y)]) | A(Q, \mathbf{Latvian})\}$
b. $\exists Q\,[A(Q, \mathbf{Latvian}) \wedge \mathbf{certain}(\exists y\,[Q(y) \wedge \mathbf{atconference}(y)])]$

A similar derivation shows that the discourse (18a,c) also satisfies the focus constraint, because Paul being certain to be at the conference (together with the assumption that Paul is a Finn) entails there being a Finn who is certain to be at the conference, and this entails the focus closure of (18c), which is (20).

(20) $\exists Q\,[A(Q, \mathbf{Latvian}) \wedge \exists y\,[Q(y) \wedge \mathbf{certain}(\mathbf{atconference}(y))]]$

As before, the same focus alternative set and focus closure result if we assume a syntactic transformaton of quantifier lowering which feeds the determination of focus semantic values. So again, also on Fox's syntactic assumptions, focus filters

representations in the required way. An appeal to scope economy is redundant.

It is not difficult to see that a narrow-scope derivation for a case where both subjects are quantfied, such as the following version of (10a), works out similarly.

(21) A Finn is certain to be at the conference, and a Latvian is, too.

The examples with a quantifier in the ellipsis VP, rather than a raising adverb, work out in a way parallel to the quantifier-lowering examples. In my canonical analysis, one assumes that the antecedent and ellipsis VPs denote higher-order properties, e.g. (22) in the case of (2). (23a,b) are representations of inverted and surface scope readings of the first sentence in (2). (23c) is the representation of the ellipsis sentence, including the focus interpretation operator, and (24) is the focus closure for this clause. Because 'some boy has the property of admiring every teacher', but not 'every teacher has the property of being admired by a boy' entails the focus closure, the discourse (23b,c) but not the discourse (23a,c) is licensed.

(22) $\lambda\mathcal{P}\forall y\,[\mathbf{teacher}(y) \rightarrow \mathcal{P}(\lambda x\mathbf{admire}(x,y))]$

(23) a. [[a boy] [admires every teacher]]$_1$ $(\forall > \exists)$

b. [[a boy] Op [admires every teacher]]$_1$ $(\exists > \forall)$

c. [Mary$_F$ Op [admires every teacher]]$\sim 1$

(24) Focus closure
$\exists x\,[\forall y\,[\mathbf{teacher}(y) \rightarrow \mathbf{admire}(x,y)]]$

## Bibliography


Fox, D.: 1995, Economy and scope, *Natural Language Semantics* 3(3), 283–341

Fox, D.: 1999, *Economy and Semantic Interpretation*, No. 35 in Linguistic Inquiry Monograph Series, MIT Press, Cambridge MA, Co-published with the MIT Working Papers in Linguistics

Hirschbühler, P.: 1982, VP deletion and across-the-board quantifier scope, in J. Pustejovsky and P. Sells (eds.), *Proceedings of the Twelfth Annual Meeting of the North Eastern Linguistic Society*, pp 132–139, U. Mass., Amherst, MA

Rooth, M.: 1992a, Ellipsis redundancy and reduction redundancy, in S. Berman and A. Hestvik (eds.), *Proceedings of the Stuttgart Ellipsis Workshop*, SFB 340, University of Stuttgart

Rooth, M.: 1992b, A theory of focus interpretation, *Natural Language Semantics* 1(1), 75–116

Rooth, M.: 1996, Focus, in S. Lappin (ed.), *Handbook of Contemporary Semantic Theory*, pp 271–297, Blackwell, Oxford

Sag, I.: 1975, *Deletion and Logical Form*, Ph.D. thesis, MIT

Schwarzschild, R.: 1999, Givenness, avoid f and other constraints on the placement of focus, *Natural Language Semantics* 7(2), 141–177

Tancredi, C.: 1992, *Deletion, Deaccenting and Presupposition*, Ph.D. thesis, MIT

# THE HELPING-EFFECT OF DATIVE CASE

ANTONIA ROTHMAYR

Department of Linguistics
University of Vienna
antonia.rothmayr@univie.ac.at

German *helfen* (help) + DAT cannot be captured by standard applicative analyses. Employing a post-Davidsonian view, the paper derives the different stative/eventive readings of *helfen*. Eventiveness is tied to DO and BECOME, but not to CAUSE. *Helfen* is related to other uses of dative in German via Brandt 2003.

One of the questions of linguistic theory is how event structure and argument structure interact with each other. This paper argues that event structure can be read off from syntactic structure directly and that the different arguments receive their interpretation because of their position within the structure.

## 1. Introduction

A number of German verbs such as *helfen* (help), *dienen* (serve), *schaden* (harm) and *nutzen* (avail) display a dative marking on one of their arguments. It has been a standard assumption in linguistic theory that these instances of case assignment are entirely lexically determined. In other words, it has been assumed that dative assignment to the object of these verbs does not follow a regular pattern. In the following, I would like to argue that there is a grammatical mechanism that determines this dative assignment. In order to do this, a look at the event structure of *helfen*-type verbs provides important insight. Hence, the lexical stipulation of dative-assignment is not necessary.

To start with, a closer look is taken at the stative/eventive ambiguity of *helfen*-type verbs along the lines of Engelberg 2005.

## 2. Kimian and Davidsonian states

In order to capture the different nature of two kinds of statives, Maienborn 2003 and Maienborn 2004 distinguish between two ontologically different kinds of event arguments. On the one hand, the well-known Davidsonian event argument is present in action sentences and in stative sentences with verbs like *sit* and *wait*. In contrast, other stative verbs like *cost* and *resemble*, as well as copular constructions contain an ontologically different argument, called the Kimian state argument.

Three diagnostics serve to distinguish between Kimian and Davidsonian stative expressions: first, manner adverbials, which specify how an event happened, can only modify a Davidsonian state but not a Kimian one. Second, event-related locative adverbials require the presence of the Davidsonian argument as well. As there is no event with Kimian states, this event cannot happen at a particular location. Third, the modifier *ein bißchen* (a little) is ambiguous between a degree reading and a temporal reading when occurring together with a Davidsonian state. In contrast, expressions that contain a Kimian state only display the degree reading.

Following Engelberg 2005, I take these three tests to diagnose the stative reading of *helfen* (help) verbs with sentential subjects.

## 3. Stative and eventive readings of *helfen*

In the next section, the claim of Engelberg 2005 that verbs like *helfen* (help) have stative readings whenever they contain a sentential subject is extended. The reading containing a Kimian state is not tied to the presence of a sentential subject. In other words, it is not the case that this interpretation arises only if the subject bears the category CP. Examples containing a non-animate DP subject are interpreted as containing a Kimian state as well. Furthermore, I will show that even the stative variant of *helfen* has a complex event structure accommodating a trigger argument, as well as the start and the gradual development of the helping-effect.

### 3.1. The active reading

The active reading of *helfen*-type verbs can be diagnosed with the help of manner adverbials such as *schnell* (quickly), event-related locative modification like *im Garten* (in the garden) and the time-span reading of *ein bißchen* (a little) which indicates that the running-time of the helpng-action was short.

(1) Die Irmi hat dem Poldi schnell ein bißchen im Garten geholfen.
The Irmi has the Poldi quickly a little in the garden helped.

### 3.2. The stative reading

Following Engelberg 2005, we see that *helfen*-verbs display a Kimian stative reading. First, manner adverbials do not modify the helping-event but the time until the helping-effect is reached.

(2) Daß ihn ein Homöopat behandelt hat, / Das Medikament hat dem Poldi
That him a homeopath treated has, / the drug has the Poldi
auf eine bekömmliche Weise geholfen.
in a salubrious way helped.

Second, event-related locative modification is not possible. If they are acceptable at all, these locative modifiers must be interpreted as frame-setters. With respect to (3),

the frame-setting interpretation would be that Poldi happened to be next to a tree, when he experienced the helping-effect.

(3) ??Daß er eine Tablette geschluckt hat, / Die Tablette hat dem Poldi neben
That he a pill swallowed has, / the pill has the Poldi next to
einem Baum geholfen.
a tree helped.

Third, *ein bißchen* (a little) can only express the degree of the helping-effect, not the temporal length of the helping-event.

(4) Daß er die Tablette geschluckt hat, / Die Tablette hat dem Poldi ein
That he the pill taken has, / the pill has the Poldi a
bißchen geholfen.
little helped.

### 3.3. The helping-effect

That a helping-effect is present in the event structure can be detected with the help of different temporal adverbials. First, time-span adverbials specify the time how long the helping-effect holds. They do not modify the help of the helping-action.

(5) Daß er eine Spritze bekommen hatte, half dem Poldi drei Stunden
That he an injection got has, helped the Poldi three hours
lang / während der Behandlung.
long / during the treatment.

Other temporal modifiers such as *after 20 minutes* express the onset of the helping-effect. They do not specify how long it takes until a helping-action starts.

(6) Die Tablette hat dem Poldi nach zwanzig Minuten geholfen.
The pill has the Poldi after zwenty minutes helped.

Finally, Engelberg 2005 notes that the helping-effect is always relativized with respect to a particular domain. In other words, the helping-effect applies to a particular domain like the financial or the health status.

(7) Das hat dem Poldi finanziell / gesundheitlich geholfen.
This has the Poldi financially / with respect to his health helped.

In sum, the tests show that even in the stative reading, there is an implicit BECOME-operator present, which expresses the onset of the helping-effect.

## 4. The structure of *helfen*

The structure of *helfen*-type verbs must, therefore, accomodate the helping-effect, the dative-argument and the actor/trigger argument. Moreover, the variability between an agent and a trigger must be captured by the underlying grammatical structure. The solution provided by Engelberg 2005 relies on non-standard semantical

principles which violate the compositionality principle. In the following, Engelberg's insights are reformulated in the framework of Brandt 2003.

### 4.1. The dative argument

Engelberg 2005 points out that the helping-effect, in order to be understood as such, must be helpful to the person denoted by the dative argument, i.e., the helping-effect must be "good" with respect to some benefactive individual. A helping-event or a helping-trigger is therefore judged as "good" if its effect is high on a personal scale. Hence, the dative argument is interpreted as a scale, to which the helping-effect is relativized.

This scalar interpretation of dative arguments has been developed by Brandt 2003 and extended by Brandt 2005 for *too*-comparatives with datives in German. According to Brandt, constructions involving a dative and a *too*-comparative as in (8) are interpreted in such a way that the degree of the comparative is evaluated with respect to a structure that is introduced by the dative argument. Thus, the book in (8) may not be too heavy in general, it might just be too heavy for this particular man.

(8) Einem Mann war das Buch zu schwer.
A man was the book to heavy.
'the book was too heavy for a man' [Brandt 2005:p.18 (65)]

Brandt argues that datives, *cipients* in his terms, are located in a special temporal projection, *t*P, which saturates a locative variable argument that is present in VP. I take the datives of *helfen*-verbs to be licensed in the specifier of *t*P as well.

In Brandt 2005, a uniform semantics for all kinds of cipients (including double object constructions, *too*-comparatives and existential constructions) is given.

(9) $\neg$AT($x_{theme}$,$p_{loc/deg}$,i) & AT($x_{theme}$,$p_{loc/deg}$,i') & i $<$ i'
[Brandt 2005:p.18 (69)]

The formula in (9) expresses that $p_{loc/deg}$, a particular (location or) degree, does not hold of $x_{theme}$ (which corresponds to the helping-effect here) at i, but it does at i'. Furthermore, Brandt argues that AT($x_{theme}$,$p_{loc/deg}$,i') is an assertion that can be understood to hold as the result of an event. On the other hand, $\neg$AT($x_{theme}$,$p_{loc/deg}$,i) corresponds to a presupposition that expresses the state of affairs before the event represented in the structure has taken place. Brandt shows that the assertion can be identified with VP, whereas the presupposition is associated with the cipient. It is possible to interpret the indices temporally, with *i* preceding *i'*. Regarding *helfen*, the degree of goodness must be higher in worlds (i') where the helping-trigger has taken place than in worlds (i) where it didn't occur.

### 4.2. The cause or trigger argument

The trigger/agent argument is licensed in the specifier of the little *v* projection (c.f. Kratzer 1996). The difference between the two can be captured with different types

of *v* (e.g. Harley 1995, Folli and Harley 2002) or with different labels (*v* vs. V; cf. Arad 1998). As it may be, I take a DO-operator to be present in *v* if there is an agent.

### 4.3. The resultant state

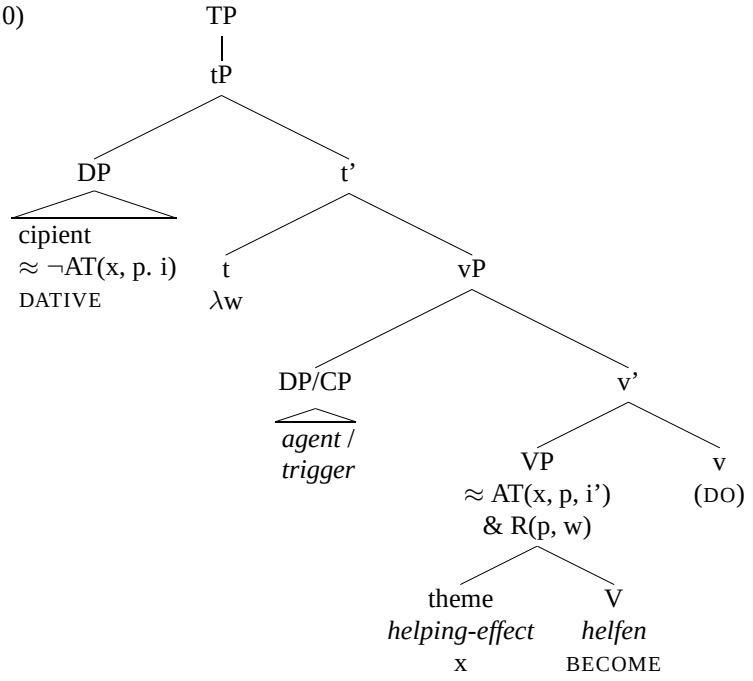
As seen above, the resultant state of the helping-action, i.e., the helping-effect, is present in the structure of *helfen*-type verbs. Following the approach by Hale and Keyser 1993, the resultant state is located in the lowest projection within the verbal structure. In particular, I take the helping-effect to be anchored in VP and to correspond to $x_{theme}$ of Brandt. The BECOME-operator, which is responsible for the gradual onset of the helping-effect, is located in the lexical entry of *help* in V.

The domain of the helping-effect is a further specification of the helping-effect itself. Thus, modifiers like *financially* specify $x_{theme}$ directly. Hence, there is no further semantic mechanism necessary to accommodate the domain of the helping-effect.

### 4.4. Summary

In sum, *helfen*-verbs have a structure as illustrated in (10). This phrase marker, modelled after Brandt 2003, consists of *v*P for the trigger/agent, *t*P for the licensing of the cipient, and VP for the resultant state/helping effect.

(10)



## 5. Conclusion

In conclusion, the paper argues for a particular event structure of *helfen*-type verbs. First, there is an agent or a trigger argument which function as the cause of the helping-effect. Second, there is a benefactive argument which carries dative case, expressing a personal scale of "goodness". Third, a helping-effect arises gradually a result of the helping-action or trigger.

The analysis of *helfen*-type verbs extends the theory of *Cipient Predication* developed by Brandt 2003, integrating the aspectual operators DO and BECOME. A CAUSE operator is not necessary to capture the behavior of *helfen*.

**Acknowledgements**

I wish to thank (in alphabetical order) Patrick Brandt, Wolfgang U. Dressler, Claudia Maienborn, Martin Prinzhorn and Edwin Williams; all remaining errors are entirely my own. This work was supported by a DOC [Doktorandenprogramm] grant of the Austrian Academy of Sciences.

## Bibliography

Arad, M.: 1998, Psych-notes, *UCL Working Papers in Linguistics* 10

Brandt, P.: 2003, *Cipient Predication: Unifying Double Object, Dative Experiencer and Existential/Presentational Constructions*, Ph.D. thesis, LOT Dissertation Series 74

Brandt, P.: 2005, *Conditioning Cipient Datives*, ms. Universität Frankfurt/Main

Engelberg, S.: 2005, Stativity, supervenience, and sentential subjects, in C. Maienborn and A. Wöllstein (eds.), *Event Arguments in Syntax, Semantics, and Discourse*, pp 45–68, Tübingen: Niemeyer

Folli, R. and Harley, H.: 2002, Consuming results in italian & english: Flavours of *v*, in P. Kempchinsky and R. Slabakova (eds.), *Aspect*, pp 1–26, Dordrecht: Kluwer

Hale, K. and Keyser, S. J.: 1993, On argument structure and the lexical expression of syntacitc relations, in K. Hale and S. J. Keyser (eds.), *The view from building 20*, pp 53–109, Cambridge, Mass.: MIT Press

Harley, H.: 1995, *Subjects, Events and Licensing*, Ph.D. thesis, MIT

Kratzer, A.: 1996, Severing the external argument from its verb, in J. Rooryck and L. Zaring (eds.), *Phrase Structure and the Lexicon*, pp 109–138, Dordrecht: Kluwer

Maienborn, C.: 2003, *Die logische Form von Kopula-Sätzen*, Berlin: Akademie Verlag

Maienborn, C.: 2004, On davidsonian and kimian states, in C. I. and K. von Heusinger (eds.), *Existence: Semantics and Syntax*, Dordrecht: Kluwer, to appear

# AGAINST PARTITIONED READINGS OF RECIPROCALS

SIVAN SABATO AND YOAD WINTER

Computer Science Department
Technion - Israel Institute of Technology
{sivansa,winter}@cs.technion.ac.il

In this paper we examine partitioned interpretations of sentences with reciprocal expressions. We study the availability of partitioned readings with definite subjects and proper name conjunctions, and show new evidence that partitioned interpretations of simple reciprocal sentences are independent of the semantics of the reciprocal expression, and are exclusively determined by the interpretation of the subject.

## 1. Partitioned Interpretations of Reciprocal Sentences

A well-known property of some sentences with reciprocal expressions is their "partitioned" interpretation. A reciprocal sentence has a partitioned interpretation if it may be evaluated as true in a situation where the antecedent set of the reciprocal is partitioned to subsets that are disjoint with respect to the predicate in the scope of the reciprocal. For example, sentence (1), from Fiengo and Lasnik 1973, has a partitioned interpretation because it is acceptable in the situation depicted in figure 1[1].

(1) The men are hitting each other.

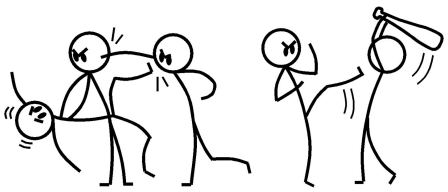


Figure 1: The men are hitting each other

[1] I am grateful to Tali Ore for creating this figure.

Another example of a partitioned interpretation is exemplified by sentence (2) below, from Dalrymple et al. 1998. This sentence is evaluated as true when there are several disjoint stacks of planks, each stack connected using the relation denoted by *stack atop*.

(2) He and scores of other inmates slept on foot-wide planks stacked atop each-other.

In the literature about reciprocals there is disagreement concerning the origin of such partitioned interpretations. The partitioning effect may result from the semantics of the reciprocal itself, but it may also follow from distribution of the antecedent set into disjoint subsets, due to independent mechanisms unrelated to reciprocal expressions.

For sentence (1), many works (see Schwarzschild 1996, Dalrymple et al. 1998, Beck 2001) agree that its partitioned interpretation arises from a partitioning of the subject NP, such that for each of the subsets Strong Reciprocity is required[2]. The operator that is assumed to create this partition in all of these works is the *cover* mechanism suggested (among others) in Schwarzschild 1996. This mechanism distributes a set denoted by a plural NP into contextually salient subsets, such that the union of the subsets equals the original set.

There is less agreement about the origin of the partitioned interpretation in other cases, as in the case of sentence (2) above. Dalrymple et al. 1998 (henceforth DKKMP) propose a system for the semantics of reciprocal expressions based on the principle they term the Strongest Meaning Hypothesis (SMH). This system includes a list of available readings, such that in each reciprocal sentence the logically strongest reading that is consistent with relevant contextual information is chosen as the interpretation of that sentence.

The partitioned interpretation of (2) is derived in DKKMP by assuming that 'unpartitioned' readings available in their system are precluded because it is impossible for 'scores' of planks to form a single stack. The SMH therefore chooses Inclusive Alternative Ordering (IAO), the weakest reading in DKKMP's system, as the one that matches this sentence. Let $A$ be the set denoted by the subject NP in the reciprocal sentence, and let $R$ be the relation denoted by the predicate in the scope of the reciprocal. Then the IAO reading of reciprocals, first proposed by Kański 1987, requires that each individual in $A$ be in a pair in $R$ with at least one other individual in $A$, as either the first or the second argument. This definition allows partitioned interpretations for sentences where the antecedent set includes more than two individuals.

Another view on this kind of example is offered in Beck 2001, where all partitioned interpretations are attributed to a general semantic process with plurals, using

[2] Strong Reciprocity requires that each of the individuals in the set be in the relation with each of the other individuals in the set.

the cover mechanism of Schwarzschild 1996. In Beck's system IAO is not generated as one of the possible meanings of reciprocal expressions.

However, the following minimal pair points to a different analysis from both DKKMP's and Beck's:

(3) The planks are stacked atop each other.

(4) Planks 1, 2, 3, and 4 are stacked atop each other.

Suppose there are two stacks of two planks each. Then sentence (3) is true although four planks could form one stack, but sentence (4) is false. This minimal pair shows that the type of the subject NP affects the availability of the partitioned interpretation: changing it from a definite plural NP to a proper name conjunction, without changing its denotation, eliminates the partitioned interpretation. The SMH mechanism cannot account for the contrast between sentences (2) and (3), in which partitioning is available, and sentence (4), in which it is not. Nor can Schwarzschild's cover mechanism, where no difference is assumed between partitioning effects of different types of plural NPs.

## 2. A New Explanation of Partitioned Interpretations

We suggest a new explanation, based on an observation made in Winter 2000, where it is shown that while definite plural NPs allow distribution to contextually salient subsets, proper name conjunction NPs resist such distribution. The following example, adapted from Winter 2000, exemplifies this contrast.

(5) The committee will commission operas to be written by teams of two composers.

   a. The composers will earn $100,000.

   b. Lloyds Webber, Penderecki, and Stockhausen will earn $100,000.

Consider a case where an opera was commissioned by the committee to be written by Lloyds Webber and Penderecki, while another opera was commissioned to be written by Lloyds Webber and Stockhausen. Each pair of composers received a total pay of $100,000 for their opera. In this situation sentence (5)a is evaluated as true, but sentence (5)b is evaluated as false. According to Winter 2000, the partitioning is available for the definite NP in (5)a because of the anaphoric power of the definite, which can combine with implicit quantification to create distribution into subsets. The same mechanism does not operate on proper name conjunctions as in (5)b since they are not anaphoric.

This analysis explains the contrast in (3)-(4), and explains in general partitioning effects with reciprocal expressions and definite antecedents. We conclude that the interpretation of the reciprocal expression itself does not allow partitioning, otherwise

this contrast would not appear. The partitioning effect in (2) follows from a similar effect of 'contextual partitioning' that also applies to plural *in*definites.

The effect of the type of the antecedent NP on the interpretation of reciprocal sentences is clearly exemplified when world knowledge allows only a partitioned interpretation. Consider the following sentences, in a situation where there are four singers:

(6) In this photo, the singers are looking into each other's eyes.

(7) #In this photo, John, Paul, George and Ringo are looking into each other's eyes.

Sentence (6) is felicitous, whereas sentence (7) is rather weird. In (7), despite world knowledge, the truth conditions derived from the reciprocal expression are not weakened to allow a partitioned interpretation.[3] If the interpretation of the reciprocal expression allowed partitioning, both sentences should have been equally felicitous. Since this is not the case, we conclude that here as well it is only a partitioning of the subject NP, available in (6) but not in (7), that allows a partitioned interpretation of the sentence.

The above contrasts suggest that there is a lower bound on the SMH-based reciprocal interpretation: it cannot be weakened to allow partitioned readings. We contend that all the cases of reciprocal sentences with partitioned interpretations are the result of an independent partitioning mechanism, while the reciprocal expression itself always has an unpartitioned interpretation.

Examining other previously suggested reciprocal interpretations that allow for partitioned interpretations shows that these partitions are indeed unrelated to the interpretation of the reciprocal expression. Sentence (8) below is brought in DKKMP as an example for One-way Weak Reciprocity (OWR), which requires that each member of the antecedent set participate in the denoted relation with another member of the antecedent set.

(8) "The captain!" said the pirates, staring at each other in surprise.

In sentence (8), OWR requires that each pirate stare at another pirate. This seems correct, as (8) is true in both figures 2(a) and 2(b). However, consider what happens when we replace the definite subject of (8) by a proper name conjunction, as in the following sentence:

(9) Morty, Charley, Oswald, Don and Bob are staring at each other.

[3]Note that a partitioned interpretation *is* available if the partition is syntactically expressed in the conjunction, as in the following variation of sentence (7) in (i) below. In this case, a partition to two pairs of singers is perfectly possible, as expected by compositionality and intersective ("Boolean") analysis of the italicized *and*.

(i) [John and Paul] *and* [George and Ringo] are looking into each other's eyes.


Figure 2: Two settings of schematic pirates

Here the truth conditions of OWR become too weak: (9) is unacceptable in figure 2(a) even though it is OK in figure 2(b). From this contrast between sentences (8) and (9) we conclude that the actual interpretation of the reciprocal expression with the predicate *stare at* requires connectivity on top of the truth conditions required by OWR. The acceptability of (8) in figure 2(a) is again attributed, as in (3) and (6), to a 'partitioning' mechanism that operates with definites, independently of the interpretation of the reciprocal expression.

To conclude, we have claimed that reciprocal expressions always require connectivity and therefore do not have partitioned interpretations. We have argued that partitioned interpretations of reciprocal sentences are always due to partitioning that is independent of the semantics of the reciprocal. In particular, partitioned interpretations of simple reciprocal sentences are exclusively dependent on the interpretation of the subject.

## Bibliography

Beck, S.: 2001, Reciprocals are definites, *Natural Language Semantics* 9, 69–138

Dalrymple, M., Kanazawa, M., Kim, Y., Mchombo, S., and Peters, S.: 1998, Reciprocal expressions and the concept of reciprocity, *Linguistics and Philosophy* 21(2), 159–210

Fiengo, R. and Lasnik, H.: 1973, The logical structure of reciprocal sentences in English, *Foundations of Langauge* 9, 447–468

Kański, Z.: 1987, Logical symmetry and natural language reciprocals, in *Proceedings of the 1987 Debrecen Symposium on Language and Logic*, pp 49–68

Schwarzschild, R.: 1996, *Pluralities*, Kluwer Academic Publishers

Winter, Y.: 2000, Distributivity and dependency, *Natural Language Semantics* 8, 27–69

# SYNTAX AND SEMANTICS OF CAUSAL *DENN* IN GERMAN

TATJANA SCHEFFLER

Department of Linguistics
University of Pennsylvania
tatjana@ling.upenn.edu

This paper presents a new analysis of *denn* (*because*) in German. In addition to causal links between propositions, *denn* can express the causation of epistemically judged propositions or of speech acts. *Denn*'s behavior is explained by two properties: On the semantics side, I show that *denn* is a conventional implicature item. Syntactically, *denn* is a coordinating conjunction of CPs. These facts explain two things. (1) Why *denn* can be used to express a wider range of causal relations than the related *weil*: *denn* can target the coerced variables over assertions as an argument, while these variables are too high for *weil*. (2) At the same time, the restrictions on the use of *denn* also follow from *denn*'s status as a coordinating conjunction and conventional implicature.

## 1. Introduction

*Weil* and *denn* are two discourse connectives in German with a (roughly) causal meaning.[1] However, they are by no means interchangeable. It has been observed in the previous literature (see e.g., Pasch et al. 2003) that German *denn* can be used in a broader range of sentences than (subordinating) *weil*. In addition to causal links between events and propositions, *denn* can express the causation of epistemically judged propositions (1a) or of speech acts (2a).

(1) a. *Es hat geregnet, denn die Straße ist ganz naß.*
b. * *Es hat geregnet, weil die Straße ganz naß ist.*
It was raining, because the street is wet.

(2) a. *Ist vom Mittag noch etwas übrig? Denn ich habe schon wieder Hunger.*
b. *?? Ist vom Mittag noch etwas übrig? Weil ich schon wieder Hunger habe.*
Is there anything left over from lunch? – Because I'm already hungry again.

[1] There has been a lot of discussion about the question whether these and similar connectives are actually *causal* (see for example Ballweg 2004). Not all sentences containing *weil* in German actually talk about causes of events or situations (i). In this paper, I will be concerned exclusively with the differences in meaning and syntax between *weil* and *denn*.

(i) Ich stehe dann morgens immer um sechs Uhr auf, weil ich dann Durchzug mache, gell.
'And then I always get up at 6 in the morning, because I air the room at that time.' (LDC: HUB)

At the same time, *denn* is not allowed in causal clauses if (i) the because-clause precedes the main clause (see 3), (ii) a direct answer to a why-question is given (4), or (iii) the content of the because-clause is evident or has been previously mentioned (see e.g., Pasch 1997).

(3) a. * *Denn es hat geregnet, ist die Straße naß.*
b. *Weil es geregnet hat, ist die Straße naß.*
Because it rained, the street is wet.

(4) a. *Warum ist die Katze gesprungen? — * Denn sie sah eine Maus.*
b. *Warum ist die Katze gesprungen? — Weil sie eine Maus sah.*
Why did the cat jump? — Because it saw a mouse.

This paper shows that *denn* is a coordinating conjunction of CPs (section 2), and semantically, a conventional implicature item (section 3). Together, these properties explain the distribution facts mentioned above.

## 2. Syntactic Properties of *denn*

*Denn*'s syntactic classification has been the subject of some discussion. While most studies mention it as a coordinating conjunction (e.g., Pasch 1997), the most recent and comprehensive study of German connectives has a different opinion. Pasch et al. 2003 treat *denn* as a special case: according to their criteria, *denn* does not subordinate (i.e., require verb-final word order in the second argument) nor embed (i.e., together with its second argument, build a constituent of the first argument). Nor, however, do they think it is coordinating.

*Denn*'s special properties can be explained even under a coordinative conjunction analysis. Unlike the other coordinative conjunctions (*und*, *oder*, etc.), *denn* can only conjoin main clauses, i.e., CPs. This explains the requirement that the conjuncts be verb-initial or verb-second. Further peculiarities of *denn* are of semantic, not syntactic, nature – I will get back to them in the following two sections. For example, *denn*-clauses can't be embedded under other functors, and both clauses that *denn* combines have to be thematic.

## 3. *Denn*'s Semantics as Conventional Implicature

Semantically, *denn* connects two events or propositions causally. *Denn*'s semantics is two-fold. Truth-conditionally, *denn* has the semantics of the logical $\wedge$. The causal meaning of *denn* is located in its conventional implicature (see Grice 1989; Karttunen and Peters 1978):

(5) In a sentence "A, *denn* B", with $[\![A]\!] = \phi$ and $[\![B]\!] = \psi$, *denn* has the following semantics:
Assertion: $\phi \wedge \psi$
Conventional Implicature: $\mathrm{CAUSE}(\psi, \phi)$

In other words, *denn* conventionally implicates that the proposition denoted by one clausal argument is caused by the proposition denoted by the other clausal argument. For *weil*, on the other hand, the causal relationship is part of the assertion.

Since *denn*'s causal meaning is a conventional implicature, the causality can not be embedded in any other functors. Bonami and Godard 2005 show tests to prove the status of a conventional implicature, in the context of evaluative adverbs in French. Their tests, applied to German *denn*, clearly show that its causal component is contributed by a conventional implicature. This contrasts with *weil*, whose causal meaning is asserted. The following data from causal clauses embedded in questions and conditionals illustrates this. Further successful tests include negation, explicit denial, embedding in the antecedent of counterfactuals, and attributions.

**Questions.** If a conventional implicature is triggered within a question, the content that is implicated cannot be understood as being in the scope of the question's illocutionary act.

(6) a. *Wer kam zu spät, weil er den Bus verpaßt hat?*
b. ?? *Wer kam zu spät, denn er hat den Bus verpaßt?*
Who was late because he missed the bus?

Example (6a) can be asked in a situation when several people were late, for (potentially) different reasons. The question is asked to clarify who of these people was late because they missed the bus (instead of for some other reason). Example (6b) cannot be used in such a situation. In fact, it is quite hard to imagine a situation that would render this sentence entirely felicitous. It seems to be possible only when it has already been established that someone was late, and that this happened because he missed the bus. If I couldn't catch the name of the person who was late, I might use (6b) to inquire this information.

**Conditionals.** Conventional implicatures cannot be embedded in the antecedent of a conditional. The following examples show that while *weil* can be embedded in a conditional, sentences with *denn*-clauses are only grammatical when the *denn*-clause is understood as a parenthetical, which stands outside of the conditional itself.

(7) a. *Wenn Peter zu spät kam, weil er den Bus verpaßt hat, war es seine eigene Schuld und er sollte bestraft werden.*
b. ** Wenn Peter zu spät kam, denn er hat den Bus verpaßt, war es seine eigene Schuld und er sollte bestraft werden.*
If Peter was late because he missed the bus, it was his own fault and he should be punished.

(8) a. *Wenn Peter zu spät kam, weil er den Bus verpaßt hat, hat er den Anfang des Films nicht gesehen.*

b. *Wenn Peter zu spät kam, denn er hat den Bus verpaßt, hat er den Anfang des Films nicht gesehen.*
If Peter was late — he missed the bus (by the way) — he won't have seen the beginning of the movie.

In examples (7–8), the consequent clauses are chosen in order to support an integrated (7) and a parenthetical (8) reading of the causal clauses, respectively. It is obvious that *denn* cannot be understood to be in the scope of the conditional. The intended meaning in example (7b) is that someone should only be punished if they were late because of their own fault (not, for example, if they were late because their car broke on the way). The sentence simply does not support this meaning.

However, *denn* is possible in the antecedent of conditionals if it is understood as a parenthetical that contributes its meaning outside of the scope of the conditional, as in (8b). Here, it is unclear whether Peter was late for the movie, but he unquestionably missed the bus (he might have taken a cab to the theater and made it in time). The *denn*-clause has the flavor of additional information that could be explicitly marked with *by the way* in English.

## 4. Consequences: Distribution of *denn* vs. *weil*

This section shows how the two facts about *denn*'s syntax and semantics explain the differences between the uses of *denn* and *weil*.

### 4.1. Three Exceptions to the Use of *denn*

The fact that *denn*-clauses cannot precede the main clause (see 3) follows straightforwardly from our elaborations about *denn*'s syntax. All coordinating conjunctions must follow their first argument.

The second exception concerns direct answers to *why*-questions, which cannot be expressed with a *denn*-clause. Note that the causal relation between the proposition in the *denn*-clause and the other proposition (expressed in the question) is presented as a conventional implicature, and not asserted. Conventional implicatures can never function as the direct answer to a question. For example, *even x, y* in English conventionally implicates that there are alternatives to x that also do y, and that x ist the most unlikely of the alternatives to do y. However, a direct question cannot be answered by these conventional implicatures (9a).

(9) a. Who is most unlikely to play the lottery? — # Even Bill plays the lottery.
b. What does being small contrast with? — # Ants are small but strong.

The third exclusion for *denn*-clauses is when the proposition in the *denn*-clause has been previously mentioned. One should take into account that truth-conditionally, *denn* means the same as *and*. Sentences where an entire conjunct of *und* (*and*) is previously mentioned are infelicitous (10). For now, it is left as an open question whether the use of *denn* is rather more constrained in this way than that of *und*.

(10) *Es wird heute regnen. —*
a. ** Ja, ich muß zuhause bleiben, denn es wird heute regnen.*
b. *?? Ja, ich muß zuhause bleiben, und es wird heute regnen.*
It's going to rain today. — Yes, I'll have to stay home, because/and it's going to rain today.

### 4.2. *Denn* in Epistemic and Speech-Act Causal Sentences

The present analysis explains why *denn* can express causation of epistemically judged propositions and speech acts while *weil* cannot. For the speech act causation (2), I adopt a proposal made for relevance conditionals (see Siegel 2005), such as (11).

(11) If you're hungry, there's pizza in the fridge.

Following Siegel, I assume that variables for potential literal acts (assertions, questions, etc.) are introduced by a meaning-shift rule when interpretation of a sentence would otherwise be divergent. Variables are introduced for the potential literal act that is commonly associated with the type of sentence, i.e., an assertion variable for declaratives, a question variable for interrogatives, etc. Obligatory existential closure applies to these variables, based on the set of relevant entities. Thus, after the meaning shift, (11) can be paraphrased as "If you're hungry, there's an assertion that pizza is in the fridge and it is relevant." Accordingly, (2a) is coerced to mean "Because I'm already hungry again, there is a relevant question whether there's anything left from lunch."

Now, since *weil* is a subordinating conjunction, the *weil*-clause is a syntactic constituent in the other argument clause. Any assertion variable that has been introduced for the matrix clause will have scope over the entire sentence, including *weil* and its clause. Therefore, *weil* cannot target this variable for scope reasons. The causal link expressed by *weil* is part of its *assertion*. Thus, the assertion variable introduced for the sentence will of course have this part in its scope.

In the case of *denn*, the *assertion* of a sentence "p, denn q" just has the content ($p \wedge q$) (which amounts to asserting p, and asserting q). In addition, both clauses can introduce potential literal acts independently, since they are complete CPs being coordinated. In this way, *denn* can target the variable introduced by the preceding clause as its argument.

For the sentences with epistemic causation like (1), the argument proceeds similarly. The epistemic operator MUST involved in these sentences is introduced by the context. *Denn* conjoins two CPs - if an epistemic MUST is inferred for the first one, sentences like (1a) are obtained.The meaning is represented in the following way:

(12) (MUST it rained ) *denn* (the street is wet)

It is a quite complicated empirical problem to resolve the data regarding *weil* entirely. Sentences like (1b) show that an epistemic MUST introduced by inference in the first argument may only have wide scope, since the sentence doesn't have the

reading that the *denn* sentence has, and is thus ungrammatical. It is yet unclear why explicit MUST sometimes can be embedded in the first argument:

(13) *? Weil sein Licht an ist, muß Peter zuhause sein.*
Because his light is on, Peter must be home.

## 5. Conclusion

This paper shows that German *denn* is a conventional implicature item, and a coordinating conjunction of CPs. Together, these facts explain why *denn* can be used to express a wider range of causal relations than the related *weil*, and why at the same time there are some restrictions on the use of *denn*.

## Acknowledgements

I am grateful to Maribel Romero for our fruitful discussions about this topic and to two anonymous reviewers for comments. Many thanks also to Ingolf Max, Beatrice Santorini, Muffy Siegel and my informants.

## Bibliography

Ballweg, J.: 2004, *Weil* – Ursachen, Gründe, Motive, in H. Blühdorn, E. Breindl, and U. H. Waßner (eds.), *Brücken schlagen. Grundlagen der Konnektorensemantik*, pp 325–332, Walter de Gruyter, Berlin

Bonami, O. and Godard, D.: 2005, Les adverbes évaluatifs dans une approche multidimensionelle du sens, in I. Choi-Jonin, M. Bras, A. Dagnac, and M. Rouquier (eds.), *Questions de Classification en Linguistique: Méthodes et Descriptions*, pp 19–37, Peter Lang, Berne

Grice, P.: 1989, *Studies in the Way of Words*, Harvard University Press, Cambridge, MA

Karttunen, L. and Peters, S.: 1978, Conventional implicature, in C.-K. Oh and D. A. Dineen (eds.), *Presupposition*, No. 11 in Syntax and Semantics, pp 1–56, Academic Press, New York

Keller, R.: 1995, The epistemic *weil*, in D. Stein and S. Wright (eds.), *Subjectivity and Subjectivisation*, pp 16–30, Cambridge University Press, Cambridge

Pasch, R.: 1997, *Weil* mit Hauptsatz – Kuckucksei im *denn*-Nest, *Deutsche Sprache* 25, 252–271

Pasch, R., Braußе, U., Breindl, E., and Waßner, U. H.: 2003, *Handbuch der deutschen Konnektoren*, Walter de Gruyter, Berlin

Siegel, M.: 2005, *Biscuit conditionals: Quantification over potential literal acts*, Ms., Temple University

# THE ROLE OF LISTS IN A CATEGORIAL ANALYSIS OF COORDINATION

MICHAEL SCHIEHLEN

Institute for Natural Language Processing
University of Stuttgart
Michael.Schiehlen@ims.uni-stuttgart.de

The paper proposes categorial analyses for coordination with multiple conjuncts, correlative coordination, and *respectively* coordination. It argues that in a categorial setting these phenomena can only be adequately analysed if a data structure of lists is introduced. To this purpose the Lambek Calculus is extended with the Kleene star, a connective that has already been explored in other substructural logics. Correspondingly, the $\lambda$ calculus is extended with list-forming operators as motivated by the analysis of the coordination phenomena.

## 1. Introduction

Like other syntactic theories, categorial grammar is concerned with the composition of form and meaning, i.e. the definition of grammatically well-formed strings and their interpretation. In categorial grammar, the Curry–Howard correspondence makes sure that the two processes are so tightly coupled that they constrain each other in a non-trivial way. Formally, the object of investigation in grammar is the set $\mathcal{L}$ of grammatically well-formed strings of words or categories. Categorial grammar provides an infinite supply of categories inductively defined from a finite set of atomic categories and two type-forming connectives, the leftward slash ($\backslash$) and the rightward slash ($/$). The compound categories can be defined in terms of the more basic categories (1); here $\cdot$ denotes string concatenation.

(1) a. $X\backslash Y = \{\, x \in \mathcal{L} : \forall y \in Y : x \cdot y \in X \,\}$

b. $X/Y = \{\, x \in \mathcal{L} : \forall y \in Y : y \cdot x \in X \,\}$

Standing for "incomplete" expressions, compound categories can be interpreted as functions: $\tau(X\backslash Y) = \tau(X/Y) = \tau(X)^{\tau(Y)}$. Another characteristic of categorial grammar is its view on syntactic derivation: Parsing is seen as a form of reasoning that can be couched in a deductive system like the Associative Lambek Calculus (ALC) (Lambek 1958). ALC provides for every type-forming connective a left rule (which eliminates the connective in one of the premises) and a right rule (which deals

with the connective in the conclusion). Right rules are associated with operations that construct new data structures (e.g. functions in ALC); left rules correspond to operations that deconstruct.

(2) a. $\dfrac{}{X : t \Rightarrow X : t}$ Id

b. $\dfrac{T \Rightarrow Y : t' \quad \Gamma[X : t(t')] \Rightarrow Z : t''}{\Gamma[X/Y : t,\ T] \Rightarrow Z : t''}$ L/ $\qquad\qquad \dfrac{T,\ Y : x \Rightarrow X : t}{T \Rightarrow X/Y : \lambda x.t}$ R/

c. $\dfrac{T \Rightarrow Y : t' \quad \Gamma[X : t(t')] \Rightarrow Z : t''}{\Gamma[T,\ X\backslash Y : t] \Rightarrow Z : t''}$ L\ $\qquad\qquad \dfrac{Y : x,\ T \Rightarrow X : t}{T \Rightarrow X\backslash Y : \lambda x.t}$ R\

This paper argues for the need of an additional type-forming connective in categorial grammar, which stands for lists. Lists of this kind have already been used in other substructural logics (Restall 2000) for the purpose of the verification of loops in programs.In the computational literature, the connective has been expressed by the Kleene star, a practice that we will follow. The paper argues that certain cases of coordination cannot be analysed in a manner compliant to the Curry–Howard correspondence unless a data structure for lists is assumed during syntactic composition.

## 2. Multi–Conjunct Coordination

The first puzzle is the ability of coordinating conjunction to take an arbitrary number of arguments. As each subcategorization frame is expressed with a different lexical entry in categorial grammar, this means that infinitely many entries are associated with a coordinating conjunction.

(3) **and**: $(((X\backslash X)\backslash X)\backslash X)/X : \lambda x_4 \lambda x_3 \lambda x_2 \lambda x_1 . x_1 \sqcap x_2 \sqcap x_3 \sqcap x_4$

To ensure a finite lexicon, a connective is required that can represent category lists of arbitrary length (also cf. (Morrill 1994, 212)): the Kleene star. In (4a), individual list items are separated by commas (just as conjuncts are in written language).

(4) $X^* = \bigcup_{n \geq 1} X^n$ where $X^1 = X$ and $X^{n+1} = \{ x_1 \cdot, \cdot x_2 : x_1 \in X, x_2 \in X^n \}$

Starred categories are interpreted as $n$-tuples of category denotations (5).

(5) $\tau(X^*) = \bigcup_{n \geq 1} \tau(X)^n$

For interpretation, we need to equip the semantic representation language (i.e. the lambda calculus) with operators for constructing and deconstructing list objects. Lists are constructed inductively with two operators (single-item lists with $\langle\!\langle . \rangle\!\rangle$, multiple-item lists with list concatenation $+$). A single item-list denotes its sole member ($[\![\langle\!\langle x \rangle\!\rangle]\!] = [\![x]\!]$); list concatenation is defined in (6).

(6) $[\![L_1 + L_2]\!] = \{ \langle x_1 \ldots x_n, y_1 \ldots y_m \rangle : \quad \langle x_1 \ldots x_n \rangle \in [\![L_1]\!] \quad \wedge$
$\langle y_1 \ldots y_m \rangle \in [\![L_2]\!] \quad \}$

List objects can be reconverted into basic objects with the **join** operation. The **join**-operator uses the binary operation $o$ to reconnect the list items. An inductive definition is given in 7. The two clauses also function as $\beta$ reduction rules.

(7) $\mathbf{join}(o, \langle\langle x \rangle\rangle) \Rightarrow x$
$\mathbf{join}(o, L_1 + L_2) \Rightarrow o(\mathbf{join}(o, L_1), \mathbf{join}(o, L_2))$

With all these operators available, we can state the lexical entry of coordinating conjunctions (8). In contrast to entry (3), the entry (8) only introduces exactly one occurrence of the conjunction connective $\sqcap$. Multiplication of this connective is carried out by $\beta$–reduction (7). It is often useful to include the last conjunct in the conjunct list as well. This can be achieved by wrapping the conjunction around the last conjunct, e.g. with the help of an additional polymorphic variable $Y$ (8).

(8) **and**: $(X\backslash(X^*/Y))/Y : \lambda t \lambda P.\mathbf{join}(\sqcap, P(t))$

In compliance with the Curry–Howard correspondence, each operator is associated with a proof rule. The two right rules **R*** and **M*** are adopted from Restall (2000, 55f).

(9) a. $\dfrac{T \Rightarrow X : x}{T \Rightarrow X^* : \langle\langle x \rangle\rangle}$ R*

b. $\dfrac{P \Rightarrow X^* : L_1 \quad Q \Rightarrow X^* : L_2}{P,\ Q \Rightarrow X^* : L_1 + L_2}$ M*

c. $\dfrac{X : a,\ X : b \Rightarrow X : c \quad \Gamma[X : \mathbf{join}(\lambda b \lambda a.c, L)] \Rightarrow Z : z}{\Gamma[X^* : L] \Rightarrow Z : z}$ L*$_1$

## 3. Partial Distribution in Multi–Conjunct Coordination

By now, we have two operators for constructing lists and one operator for deconstructing lists. A fourth operator can be used to apply functions to lists directly. The **map**-operator, inductively defined in (10), modifies a list $L$ by applying a function $f$ to each list item. Again the two clauses also function as $\beta$ reduction rules.

(10) $\mathbf{map}(f, \langle\langle x \rangle\rangle) \Rightarrow \langle\langle f(x) \rangle\rangle$
$\mathbf{map}(f, L_1 + L_2) \Rightarrow \mathbf{map}(f, L_1) + \mathbf{map}(f, L_2)$

The proof rule for the **map** operator, again a left rule, is given in (11).

(11) $\dfrac{\Gamma[X : x] \Rightarrow Y : y}{\Gamma[X^* : L] \Rightarrow Y^* : \mathbf{map}(\lambda x.y, L)}$ L*$_2$

In an analysis of coordination, the **map** operator can be used to distribute over conjuncts. In cases that require such distribution, a functor $f$ which syntactically applies to a coordination semantically applies to each of the conjuncts. The idea of the analysis is as follows: The list types allow to separate the task of collecting the conjuncts (rules **R*** and **M***) from the application of the conjunction functor (rule **L*** $_1$). Hence a scopal item can get scope over all conjuncts without leaving the scope of conjunction: by applying to the list (rule **L*** $_2$).

A syntactic account of distribution has certain advantages with respect to effects that concern the interplay of syntax and semantics. Larson (1985, 220) has observed that the placement of the correlate of a coordinating conjunction fixes the conjunction's semantic scope. So, the position of the correlate in example (12) enforces wide scope for the disjunction. The effect follows if we assume that the correlate is picked up as the last argument of the conjunction (so that e.g. *or* would get the entry $((X\backslash \textit{either})\backslash(X^*/Y))/Y$).

(12) Mary is either looking for a maid or a cook.

Hudson (1989, 89) discusses cases where a functor is only partially distributed, i.e. distributed not over the entire conjunct list, but only over a contiguous sublist (13).

(13) a. in the United States, (the) Netherlands and in England

b. either in England, in the United States or (the) Netherlands

Cases of partial distribution provide strong motivation for the assumption that list formation and conjunct interpretation should be separated. Only such a factorization allows the propagation of the conjunction connective to sublists that do not include the coordinating conjunction. To analyse partial distribution, a list must be decomposable into arbitrary sublists. Hence, a general operation of list composition (as in rule **M***) is required; it would not suffice to only consider lists of a string-like structure, i.e. lists where a always a single element is prepended or appended.

## 4. Respectively–Coordination

A third argument for a process of list formation is provided by *respectively*–coordination (ResC). In ResC several surface conjunctions are conflated to a single functor on the semantic form (14).

(14) John and Peter love Mary and hate Sue, respectively.
love(j,m) $\wedge$ hate(p,s)

An occurrence of ResC consists of at least two coordinations. All but one coordination are modified by the adverb *respectively*. We will call the unique coordination without *respectively* the governing member, and all other coordinations the dependent members. Occurrences of ResC with more than two members (cf. (15) from (Schachter 1973, 390)) can be regarded as recursive applications of binary ResC.

(15) [John and Bill went to New York and Chicago respectively] on Monday and Wednesday respectively.

Each member coordination must have exactly the same number of conjuncts (or, as we shall say, the same arity). In the interpretation, the conjuncts of the members are correlated so that every $i$-th conjunct relates to the other $i$-th conjuncts. The correlating behaviour is reminiscent of the scalar product (16).

(16) $\langle x_1, \ldots, x_n \rangle \cdot \langle y_1, \ldots, y_n \rangle = x_1 y_1 + \ldots + x_n y_n$

The scalar product will be taken as a guideline in developping a theory of ResC. The fact that ResC can be stacked (cf. example (15)) makes necessary an adjustment, however. The linguistic operation, which will be called **vecp** for vector product, transforms its two argument lists not into a basic object but into yet another list. The operation is inductively be defined in (17). Again the definition clauses of (17) also serve as $\beta$–reduction rules.

(17) $\mathbf{vecp}(f, \langle\!\langle x_1 \rangle\!\rangle, , \langle\!\langle x_2 \rangle\!\rangle) \Rightarrow \langle\!\langle f(x_1)(x_2) \rangle\!\rangle$
$\mathbf{vecp}(f, \langle\!\langle x_1 \rangle\!\rangle + L_1, \langle\!\langle x_2 \rangle\!\rangle + L_2) \Rightarrow \langle\!\langle f(x_1)(x_2) \rangle\!\rangle + \mathbf{vecp}(f, L_1, L_2)$
$\mathbf{vecp}(f, L_1, (L_2 + L_3) + L_4) \Rightarrow \mathbf{vecp}(f, L_1, L_2 + (L_3 + L_4))$
$\mathbf{vecp}(f, (L_1 + L_2) + L_3, L_4) \Rightarrow \mathbf{vecp}(f, L_1 + (L_2 + L_3), L_4)$

**vecp** is the only list operator for which order is relevant. But for **vecp** and ResC, multisets could have been used instead of lists. By the Curry–Howard correspondence, **vecp** can be correlated with the following left rule (18).

(18) $$\frac{X : x,\ Y : y \Rightarrow U : u \quad \Gamma[U^* : \mathbf{vecp}(\lambda x \lambda y.u, L_1, L_2)] \Rightarrow Z}{\Gamma[X^* : L_1,\ Y^* : L_2] \Rightarrow Z}\ \mathrm{L}^*{}_3$$

The analysis has to come to grips with the fact that in the final representation the coordinating conjunctions of all member coordinations are conflated to just one instance. We assume that this instance is triggered by the governing member. Conjunctions in dependent coordinations merely pass on their conjunct lists (19).

(19) **and** $((X^* \mid \textit{respectively}) \backslash (X^* / Y)) / Y : \lambda t \lambda P \lambda r.P(t)$

The conjunction in the governing coordination has a local effect (essentially null) and a global effect (introducing its meaning). In this respect it is similar to quantifiers, and the techniques used for quantifier raising may be applied (e.g. wrapping, polymorphism, or a lexical entry communicating with a unary rule).

(20) **and** locally: $(X^* \backslash (X^* / Y)) / Y : \lambda t \lambda P.P(t)$
globally: $X \mid X^* : \lambda L.\mathbf{join}(\sqcap, L)$

We assume that some modality or feature mechanism controls the communication between the rules and lexical entries in (18, 19, 20) and ensures e.g. the presence of exactly one governing coordination and at least two members.

The consequences of ResC for the complexity of syntax have been hotly debated in the literature. ResC has been used as an argument that natural language is not even mildly context-sensitive (Kac 1987). To salvage context-freeness, it is important that the same–arity restriction be tested not before in semantics. That semantics plays an important role is obvious in constructions with plural noun phrases (21): The number of partition cells, though relevant for ResC, has no reflex in syntax (Pullum and Gazdar 1982, 500.fn(10)).

(21) during the period of squabbling between court factions supporting Russia or Japan respectively (BNC)

In the analysis presented here, the arity restriction is not checked during composition but rather during $\beta$–reduction (17). For invalid sentences, the parser produces a result, but the result includes unreduced lists.

## 5. Conclusion

The paper has argued for the necessity of including lists in the categorial analysis of coordination. First, without lists, either the phrase structure rule base or the categorial lexicon will cease to be finite, as a single coordinating conjunction can connect arbitrarily many conjuncts. Second, lists allow for an explicit modelling of distribution over conjuncts. Such a treatment is needed e.g. to adequately express the mutual restrictions between syntax and options for distribution. It is also needed to account for partial distribution, i.e. distribution over only a subset of conjuncts in a multi–conjunct coordination. Finally, lists are a handy tool in analysing *respectively* coordination. I would like to thank Hans Kamp, Kristina Spranger, and an anonymous reviewer for their helpful comments.

## Bibliography

Hudson, R. A.: 1989, Gapping and grammatical relations, *Journal of Linguistics* 25

Kac, M. B.: 1987, Surface Transitivity, *Respectively* Coordination and Context-Freeness, *Natural Language and Linguistic Theory* 5, 441–452

Lambek, J.: 1958, The mathematics of sentence structure, *American Mathematical Monthly* 65, 154–170,

Larson, R. K.: 1985, On the Syntax of Disjunction Scope, *Natural Language and Linguistic Theory* 3, 217–265

Morrill, G. V.: 1994, *Type Logical Grammar*, Kluwer

Pullum, G. and Gazdar, G.: 1982, Natural Languages and Context-Free Languages, *Linguistics and Philosophy* 4, 471–504

Restall, G.: 2000, *An Introduction to Substructural Logics*, Routledge

Schachter, P.: 1973, Conjunction, in R. P. Stockwell, P. Schachter, and B. H. Partee (eds.), *The Major Syntactic Structures of English*, pp 294–418, Holt, Rinehart & Winston

# Transparency: An Incremental Account of Presupposition Projection

Philippe Schlenker
UCLA & IJN

We sketch a theory in which presuppositions do not directly impose conditions on the *context set,* but rather on the *contextual meaning* of a sentence. Specifically, a part of an expression's meaning which is marked as presupposed should satisfy a principle of *Transparency*, according to which *this part can be disregarded without affecting the contextual meaning of the sentence*. We argue that if *Transparency* is checked incrementally, i.e. as soon as a clause is pronounced, it yields a predictive account of presupposition projection: unlike competing theories, it *derives the projection behavior of connectives from their bivalent semantic contribution*. We speculate that *Transparency* originates from a more general pragmatic principle, *Be Articulate!*, which states that *one should not say too much at the same time*, i.e. express a meaning that is too complex with a single expression. *Transparency* is a way to satisfy *Be Articulate!* even when an expression with a complex meaning is uttered because it ensures that part of this meaning can be disregarded.

## 1. Programmatic Outline: *Be Articulate!* and *Transparency*

Two main questions can be asked about presuppositions: (i) How are they triggered? (ii) How are they projected? In ground-breaking work, Heim 1983 gave a *lexical* answer to both questions (similar remarks apply to the DRT accounts of van der Sandt 1993 and Geurts 1999; these raise several problems for Heim's theory, which we inherit). Heim made the following assumptions:

**(i)** Presuppositions are triggered lexically, in a context-insensitive fashion (though *accommodation* is context-sensitive). Thus whenever *John knows that p* is uttered, p will be triggered as a presupposition.

**(ii)** The projection behavior of connectives and operators is encoded in their lexical entry. For example, if C is a context set, it is stipulated that C[F and G]=(C[F])[G]. [For an atomic proposition *F*, we write *F* as $\underline{p}p'$ if F contributes a presupposition p and an assertion p'. In this case, Heim's theory specifies that the update of C with F is C[F]=C[$\underline{p}$p']=# iff for some w∈C, p(w)≠1. If ≠#, C[F]=C[$\underline{p}$p']={w∈C: p(w)=1}].

Assumptions (i) and (ii) both raise the same explanatory problem:

**(i')** Why could there not be a verb *know**, which has the same global (i.e. assertive + presuppositional) content as *know*, but a different presupposition? For instance we could imagine that *John knows* that p* has no presupposition but asserts that *p and John believes that p*. We rarely encounter words such as *know** (but see Abusch 2002 for a different opinion). Why? Heim 1983 provides no answer.

**(ii')** Why could there not be a word *and** which had the same logical contribution as *and* but a different projection behavior? For instance we could imagine that C[p and* q]=(C[q])[p]. But there seems to be no word such as *and**. Why? Here too Heim 1983 (criticized by Soames 1989 and Heim 1992) gives no answer.

We will sketch a purely pragmatic account of presupposition triggering and presupposition projection. Our attempt is purely programmatic with respect to the Triggering Problem. On the other hand we offer a precise algorithm for the Projection Problem, one that is *predictive*, in the sense that it can be applied to any connective as soon as its bivalent semantic contribution is known (neither Heim 1983 nor DRT are predictive in this sense). Our account is stated within a fully classical (bivalent) framework, and it has the following structure (the two parts could well be separated, but the result would be conceptually less natural):

**1.** A general pragmatic principle, ***Be Articulate!*****,** specifies that *one should not say too much at the same time*, in the sense that one should not express a complex meaning with a single expression. For instance, in *John fell*, the single word *fell* contributes -very roughly- the information that i) John was standing up (=*u* for short), ii) he underwent an involuntary motion (=*i*), and iii) He ended up lying down (=*d*). Unless one is explicitly interested in this complex conjunction *uid*, *Be Articulate* risks being violated: it would be better practice to articulate the intended meaning in separate parts, e.g. as *John was standing but he fell*.

**2.** In case *Be Articulate!* might seem to be violated, there is a way to salvage the principle by assuming that part of the meaning of the offending term is *transparent*, in the sense that one can *disregard it without changing the contextual meaning of what is said.* Specifically, ***Transparency*** states that one can *erase* one of the conjuncts that make up the meaning of the offending term, and still obtain a sentence which, given the assumptions of the conversation, is equivalent to the original one. *Transparency* is checked in two steps:
a. *Selection* (=Triggering Problem)*:* First, one divides up the meaning of the offending term into two parts, and chooses (on pragmatic grounds to be determined) which one should be transparent; we write this part as underlined. How this selection process is performed may depend on the context. For instance, *He didn't fall* typically presupposes that the agent was standing up; we have in this case an analysis of the sentence as *not f = not(uid)*. But one may utter the sentence felicitously to reassure a concerned mother who just saw that her son is lying on the floor crying (*Don't worry, he didn't fall*); here one seems to be presupposing that the little boy is lying on the floor, not that he was standing up right before [*not f = not(uid)*]. To see a third kind of situation, suppose that we saw someone come off a cliff. If I say that *he didn't fall*, I am probably presupposing that he was standing up and is now lying down, and denying that the motion was involuntary. Thus depending on the context, *John didn't fall* may be variably analyzed as *not uid, not uid, not uid* (or as *uid* if one is explicitly interested in the conjunction).
b. *Incremental Verification* (=Projection Problem)*:* Second, as soon as the offending clause is pronounced, one checks that, no matter what the end of the sentence will be, and no matter what the semantic content of the non-underlined part of the clause might be, *Transparency* will be satisfied. Suppose that *John didn't fall* is analyzed as *not uid*. In a context set C, we want to ensure that no matter what the end β of the sentence is:
(1) C |= ∀X [not(uX)β ↔ not(X)β)
This will turn out to require that C |= u, as is desired. By contrast, if the sentence uttered is *John was standing and he didn't fall*, it can be understood with no presupposition whatsoever, because *Transparency* is automatically satisfied - no matter what C is, we know that:
(2) C |= ∀X [(u and not(uX))e ↔ (u and not(X) e)]
We now develop in greater detail our account of the Projection Problem. Programmatic remarks on the Triggering Problem are included in the last section.

**2. The Projection Problem I: Principles**

**A) The Stalnaker/Heim Dilemma:** Unlike the standard Stalnaker/Heim account of presupposition projection, the present analysis does not take presuppositions to be constraints on the *context*, but rather on the *contextual meaning* of a sentence (specifically: a presupposition is a part of the meaning of a clause that one should be allowed to disregard without changing the truth conditions of the sentence). We take the Stalnaker/Heim analysis to have the following logic, which leads straight into a dilemma.
**1.** *Assumption:* When a clause *pp'* with presupposition p is uttered, it requires that p be taken for granted in the context (i.e. context set) of utterance .
**2.** *Observation:* In some cases, the *Assumption* seems to be violated, e.g. in *It is raining and John knows that it is*, which does not presuppose anything.
**3.** *Conclusion*: The notion of 'context' must be ramified. In the course of the evaluation of a sentence, there are a variety of *local contexts*, which are obtained as modifications of the initial context. In *It is raining and John knows that it is*, the local context obtained after the first conjunct is evaluated is one in which the presupposition that it is raining is indeed satisfied.

Stalnaker and Heim differ in the way in which they set up the theory of local contexts. 1) In the case of conjunction, Stalnaker 1974 argues that presupposition projection can be explained in pragmatic terms: in *p and q*, *q* is evaluated in the initial context set *as modified by the assertion of p*. This is a plausible analysis, but only because a conjunction can be seen as a succession of two assertions. The account does not extend to other connectives, such as disjunction (*p or q* can certainly not be analyzed as a succession of assertions). 2) Heim 1983 abandons Stalnaker's pragmatic explanation, and posits that the

way in which a connective modifies the context set is stipulated in its lexical entry. This account can extended to any connective, but it fails to be explanatory: it does not explain why the conjunction we find in natural language is *and* rather than *and**. The dilemma is thus between a pragmatic - and explanatory - analysis that works for conjunction but not for all other connectives; and a lexical account that works for all connectives but is not explanatory. We conclude that a different course should be taken: the *Observation* should be seen as *refuting the Assumption*. Presuppositions do not directly impose something on the *context*, but rather on the (contextual) *meaning* of a sentence.

**B) Transparency:** Let us assume that Selection has been performed, and that we are given a Logical Form that includes elements of the form $\underline{p}p'$, where $\underline{p}$ is underlined to indicate that it must be transparent. The principle to be satisfied is stated in (3).

(3) *Transparency:* For any initial part $\alpha\ \underline{p}p'$ of a sentence uttered in a background of assumptions C, where $\underline{p}p'$ is the semantic composition of an IP (=a clause), it should be the case that for any sentence completion β, C|= $\forall X(\alpha\ (\underline{p}X)\ \beta \Leftrightarrow \alpha\ X\ \beta)$, where *X* is a propositional variable[1].

We will assume that *Transparency* is checked incrementally with respect to *linear order*. There are variants of theory, however, in which *Transparency* is checked following (a) order of processing (whatever this turns out to be), or (b) order c-command (scope) in a top-down system. We do not explore these alternatives here[2]. Instead, we immediately turn to examples and show that in simple cases the *Transparency* theory can emulate the results of Heim 1983.

**3. The Projection Problem II: Examples**

**A) Connectives I - Standard Cases (*not, and, if*):** We start with an extremely simple syntax for the object language[3]:

(4) *Syntax:* F ::= p | (pp') | not F | [F and F'] | [F or F'] | [if F]F'

Boolean connectives have their standard semantics, and *if F, G* is taken to be a strict implication: with background assumptions C, *if F, G* evaluated in any C-world is true iff every C-world that satisfies *F* also satisfies *G*. As before, $\underline{p}p'$ is interpreted as a simple conjunction (the syntax is intended to indicate that the conjuncts correspond to a single lexical item). We go through some representative examples, stating in each case what *Transparency* requires and how it derives the correct projection behavior.

**Example 1. Sentences starting with *(pp')* [e.g. *(pp') and ..., (pp') or ...* ]**

*Transparency* requires that for any sentence completion β, C |= $\forall X\ ((\underline{p}X)\beta \Leftrightarrow X\ \beta)$

*Claim: Transparency* is satisfied iff C |= p.

**i.** If C |=p, C |= $\forall X\ ((\underline{p}X) \Leftrightarrow p')$, and the result follows. **ii.** Taking β to be the null string, *Transparency* requires that C |= $\forall X\ ((\underline{p}X) \Leftrightarrow X)$. Taking X to be a tautology, we obtain that C |= p.

**Example 2. Sentences starting with *[not (pp')]***

*Transparency* requires that for any sentence completion β, C |= $\forall X\ ([\text{not}(\underline{p}X)]\beta \Leftrightarrow [\text{not } X]\beta)$.

*Claim: Transparency* is satisfied iff C |= p.

---

[1]The quantification over X will have to be eliminated or justified in future work. It is an unpleasant feature of the present analysis. A more natural condition would be: C |= $(\alpha\ (\underline{p}p')\ \beta \Leftrightarrow \alpha\ p'\ \beta)$, but this is not quite strong enough to derive the desired results.

[2]Other variants of the system would be obtained if instead of requiring that the 'stripped' version of the sentence is equivalent to the original one, one simply stipulated that there should be no asymmetric entailment between the two. In monotonic environments, this will make the same predictions as the present theory. But for non-monotonic environments, the results will be different (the prediction is that non-monotonic operators should 'filter out' presuppositions, which sometimes appears to be the case).

[3]When we discuss the effects of Transparency, we enrich this language with propositional variables, quantifiers, material implication (⇒), and material equivalence (⇔); and we allow $\underline{p}$ to be an atom.

**i.** If C |=$\underline{p}$, the result follows immediately. **ii.** Taking β to be the null string, we have C |= ∀X (not($\underline{p}$X) ⇔ not X), hence C |= ∀X (($\underline{p}$X) ⇔ X), and thus from Example 1, ii: C |= p.

**Example 3. Sentences starting with *[p and ($\underline{q}$q')].***

*Transparency* requires that for any sentence completion β, C |= ∀X ([p and ($\underline{q}$X)] β ⇔ [p and X] β)

*Claim: Transparency* is satisfied iff C |= p ⇒ $\underline{q}$.

**i.** If C |= p ⇒ $\underline{q}$, C |= ∀X ([p and ($\underline{q}$X)] ⇔ [p and X]) and hence for any sentence completion β, C |= ∀q' ([p and ($\underline{q}$q')]β ⇔ [p and q'] β) **ii.** By taking β to be the null string and X to be a tautology, *Transparency* entails that C |= p ⇒ [p and $\underline{q}$], and hence that C |= p ⇒ $\underline{q}$.

**Example 4. Sentences starting with *[if ($\underline{p}$p')]***

We start by observing that, given our semantics, if w is a C-world, w |= [if F] F' if and only if C |= F ⇒ F'.

*Transparency* requires that for all sentence completions β, C |= ∀X ([if ($\underline{p}$X)]β ⇔ [if X] β)

*Claim: Transparency* is satisfied iff C |= $\underline{p}$.

**i.** Clearly, if C |= $\underline{p}$, C |= ∀X (($\underline{p}$X) ⇔ X), from which it follows that C |= ∀X ([if ($\underline{p}$X)]β ⇔ [if X]β). **ii.** Taking β to be $\underline{p}$ and X to be a tautology, *Transparency* entails in particular that C |= ([if ($\underline{p}$X)]$\underline{p}$ ⇔ [if X]$\underline{p}$). The left-hand side is tautology, and thus C |= $\underline{p}$.

**Example 5. Sentences starting with *[if p]($\underline{q}$q')***

*Transparency* requires that for all sentence completions β, C |= ∀X ([if p]($\underline{q}$X) β ⇔ [if p]X β)

*Claim: Transparency* is satisfied iff C |= p ⇒ $\underline{q}$

**i.** If C |= p ⇒ $\underline{q}$, C |= ∀X ([if p]($\underline{q}$X) ⇔ [if p]X), from which it follows that for any sentence completion β, C |= ∀X ([if p]($\underline{q}$X)] β ⇔ [if p] X β)[4]. **ii.** Taking β to be empty and X to be a tautology, *Transparency* entails that C |= [if p] $\underline{q}$ ⇔ [if p]X. The right-hand side is a tautology, and thus C |= p ⇒ $\underline{q}$.[5]

**B) Connectives II - Other Cases (*or, unless, while):*** Heim 1983 does not discuss the projection behavior of disjunctions. There are a variety of positions in the literature (see Krahmer 1998, Beaver 2001). Following Beaver 2001, we take the correct result to be that *$\underline{p}$p' or q* presupposes that *p*, and *p or $\underline{q}$q'* presupposes that *if not p, q*. Whatever their stand on this issue, competing theories must stipulate the projection behavior of *or,* which does not follow from anything else. By contrast, our algorithm makes precise predictions: as shown in Example 1, *$\underline{p}$p' or q* presupposes *p*; and *p or $\underline{q}$q'* presupposes *if not p, q:*

---

[4]The result follows because our very simple syntax guarantees that any sentence starting with *[if F]G* has *[if F]G* as a syntactic unit (otherwise *G* should be preceded by: *[* ).

[5]An *Amsterdam Colloquium* reviewer observes that post-posed *if-clauses* might well have the same projection behavior as pre-posed ones: $^{Ok}$*If there is a reviewer, the reviewer is mad,* $^{Ok}$*The reviewer is mad, if there is a reviewer.* Is this a problem? Given our statement of *Transparency*, it all depends what the syntactic position of the post-posed *if*-clause is. If it is not possible to have a complete IP without including it, then we predict that the projection behavior of post-posed *if*-clauses should indeed be identical to that of pre-posed *if*-clauses. Now Bhatt & Pancheva 2001 argue that post-posed are attached quite low, as suggested by the Condition C effect that obtains in $^{\#}$*She$_i$ yells at Bill if Mary$_i$ is angry*[5]. The results are arguably similar in $^{\#}$*He$_i$ is mad, if [the reviewer]$_i$ exists* (which contrasts with *If he$_i$ exists, [the reviewer]$_i$ is mad*). This suggests that the post-posed *if*-clause is in the scope of the matrix subject, and hence belongs to the smallest IP that includes the matrix verb. As a result, only after the post-posed *if*-clause is processed can *Transparency* be checked, which predicts that *The reviewer is mad, if there is a reviewer* should indeed be acceptable...

[7]Are these predictions correct? *This house has no bathroom or the bathroom is in a funny place* (after Partee) suggests that the analysis of Example 6 might be right. But arguably *The bathroom is in a funny place or this house has no bathroom* is also acceptable, which does not square well with the claim that *($\underline{p}$p' or q)* presupposes that p. But as pointed out by B. Spector (p.c.), there are other cases that suggest that there is a systematic asymmetry in the projection behavior of disjunctions, as illustrated by the following contrast:

**Example 6. Sentences starting with *[p or (q̲q')]*.**
*Transparency* requires that for any sentence completion β, C |= ∀X ([p or (q̲X)] β ⇔ [p or X] β).
*Claim: Transparency* is satisfied iff C |= (not p) ⇒ q̲
**i.** If C |= (not p) ⇒ q̲, C |= ∀X ([p or (q̲X)] ⇔ [p or X]) [this follows from the propositional logic equivalence between *p or r* and *p or ((not p) and r)*]. Thus for any sentence completion β, C |= ∀X ([p and (q̲X)]β ⇔ [p and X] β) **ii.** By taking β to be the null string and X to be a tautology, *Transparency* entails that C |= [p or q̲] ⇔ [p or X]. The right-hand side is tautology, thus C|= [p or q̲], i.e. C|= (not p) ⇒ q̲[7].

Heim 1983 makes no predictions about other connectives that she does not consider, such as *unless* or *while*. But the present analysis is more constrained. From the equivalence between *Unless F, G* and *Unless F, (not F) and G*, we predict that any presupposition of G entailed by *not F* should automatically be transparent. This prediction is borne out in (5), which presupposes nothing at all:
(5) Unless John didn't come, Mary will know that he is here (*presupposes nothing)*
(5) has the form *unless F, q̲q'*, where *not F* entails *q̲* (specifically: *not (John didn't come)* entails: *John is here*). This accounts for the data. Turning now to *while*, the equivalence between *While F, G* and *While F, F and G* explains the facts in (6):
(6) While John was working for the KGB, Mary knew that he wasn't truthful about his professional situation.
(6) is of the form *While F, q̲q'*, where *F* contextually entails *q* (because a spy isn't truthful about his professional situation). *Transparency* is automatically satisfied.
**C) Extension - Quantifiers (Simple Cases):** Presupposition projection in quantified structures is a notoriously hairy topic, which we only treat superficially by considering *[every P](Q̲Q')*, *[at least one P](Q̲Q')* and *[no P](Q̲Q')* (we extend out notation from propositional letters to predicates: the underlined part is the presuppositional one, and concatenation is interpreted as generalized conjunction). In all three cases Heim 1983 predicts the same presupposition, namely that *every P-individual satisfies Q*. For better or worse, we match this result.
**Example 7. Sentences starting with *[Every P] (Q̲Q')***
*Transparency* requires that for any sentence completion β, C |= ∀Y ([Every P] Q̲Y β ⇔ [Every P]Y β), where Y is a predicate variable.
*Claim: Transparency* is satisfied iff C |= [Every P] Q̲
**i.** Clearly, *Transparency* is satisfied if every P-individual is a Q̲-individual. **ii.** If some P-individual, say i, is not a Q̲-individual, *Transparency* fails: take β to be the null string, and take Y to hold of every individual. Then the right-hand side holds, but the left-hand side doesn't.
**Example 8. Sentences starting with *[At least one P] (Q̲Q')***
*Transparency* requires that for any sentence completion β, C |= ∀Y([At least one P] *(Q̲Y)* β ⇔ [At least one P] Y β)
*Claim: Transparency* is satisfied iff C |= [Every P]Q̲
**i.** Clearly, *Transparency* is satisfied if every P-individual is a Q̲-individual. **ii.** If some P-individual, say i, is not a Q̲-individual, *Transparency* fails: take β to be the null string, and take Y to hold only of i. Then the right-hand side holds, but the left-hand side doesn't. [These predictions are notoriously too strong for indefinites, as in *A fat man was pushing his bicycle*. We leave this for the future...]
**Example 9. Sentences starting with *[No P] (Q̲Q')***
*Transparency* requires that for any sentence completion β, C |= ∀Y([No P] (Q̲Y) β ⇔ [No P] Y β)
*Claim: Transparency* is satisfied iff C |= [Every P]Q̲
**i.** Clearly, *Transparency* is satisfied if every P-individual is a Q̲-individual. **ii.** If some P-individual, say i, is not a Q̲-individual, *Transparency* fails: take β to be the null string, and take Y to hold only of i. The left-hand side is true (since i is not a Q̲-individual, Q̲Y has an empty extension), but the right-hand side is false (it is refuted by i itself).

**4. The Triggering Problem: Remarks**

Our projection algorithm could in principle be adapted to any solution to the Triggering Problem. As long as there is a way to determine which elements are 'underlined', *Transparency* can be applied to yield the predictions we have laid out. However one would then like to know *why Transparency* should hold in the first place. In the present framework, *Transparency* can be seen as a strategy of *complexity reduction*, which guarantees that *Be Articulate!* is satisfied even when a complex meaning is expressed by a single expression, thanks to the assumption that part of this meaning is eliminable. We now present preliminary evidence in favor of this pragmatic analysis as it applies to the Triggering Problem.

**1)** As was observed at the outset, in some cases, such as *fall*, the generalization appears to simply be that *some* part of the meaning should be presupposed; which one it is would seem to be context-dependent[8]. In more recalcitrant cases, similar effects can be obtained by manipulating the preceding discourse. *Do you know that Mary is pregnant* normally presupposes that Mary is pregnant. By contrast, *Do you know that Mary is pregnant or do you believe it?* presupposes that the addressee believes that Mary is pregnant.

**2)** As pointed out by Simons 2001, presuppositional effects can be obtained through adverbial modification: *None of my students arrived on time* typically implies that each of my students arrived. It is difficult to see how this could come from an implicature (with a scale <*arrive, arrive on time*>), as this analysis would predict a weaker inference, namely that some of my students arrived[9]. From the present perspective, the facts follow because *arrived on time* (by contrast with *arrived, and did so on time*) violates *Be Articulate!* unless part of the meaning is taken to be transparent. This triggers a presuppositional phenomenon, as desired.

**3)** Obviously these remarks only scratch the surface of the Triggering Problem. At best they indicate that presuppositions are triggered when an expression would otherwise violate *Be Articulate!*. But this leaves entirely open the issue of *which part of the meaning is selected to be transparent*. This we leave for future research.

**Acknowledgements**

Thanks to D. Fox, B. Geurts, B. Spector, the audience of *Semantiknetzwerk*, and a reviewer for helpful remarks. The author gratefully acknowledges the financial support of the ACLS (*Ryskamp Fellowship*).

**Bibliography**

**Abusch, D.**: 2002, Lexical Alternatives as a Source of Pragmatic Presuppositions; **Beaver, D.** 2001. *Presupposition and Assertion in Dynamic Semantics*; **Bhatt, R. & Pancheva, R.**: 2001, The Syntax of Conditionals. **Geurts, B.** 1999. *Presupposition and Pronouns*. **Heim, I.:** 1983, On the Projection Problem for Presuppositions. **Heim, I.**: 1992, Presupposition Projection and the Semantics of Attitude Verbs; **Krahmer, E.**: 1998, *Presupposition and Anaphora*; **Simons, M.**: 2001, On the Conversational Basis of some Presuppositions; **Soames, S.**: 1989, Presupposition; **Stalnaker, R.**: 1974, Pragmatic Presuppositions; **van der Sandt, R.**: 1993, Presupposition Projection as Anaphora Resolution.

[8]The effect can be reproduced in other cases. *None of these kids is starting to smoke* typically presupposes that each of the kids was heretofore a non-smoker. But if I utter the sentence while watching some teenagers smoking, we obtain the opposite pattern: it is then presupposed that each of kids is a smoker, and it is asserted that each of them was one before.

[9]An alternative possibility, however, is that the standard theory of scalar implicatures is incorrect, as has been recently argued by Chierchia and others. It is thus important to perform a direct comparison between adverbial modification and clear instances of scalar implicatures in similar environments. Focusing on the scale <*or, and*>, we can consider *None of my students read Chomsky and Montague*. The question is whether this simply implies that some of my students read Chomsky or Montague, or that *each* of my students read Chomsky or Montague. The data are subtle. My impression is that without focus on *and* we only get the weak implication. With focus on *and*, I believe the strong implication is obtained. I leave this issue for future research.

# EXHAUSTIVE IMPERATIVES

MAGDALENA SCHWAGER

Institut für Kognitive Linguistik
Johann Wolfgang Goethe-Universtität Frankfurt am Main
magdalena@schwager.at

**Abstract.** Usually imperatives show a tight link to necessity, but examples modified by *for example* provide evidence for possibility as their semantic core. It is argued that the possibility operator is normally turned into necessity by a covert exhaustifier whose application can be blocked by overt *for example*.

## 1. Introduction: a puzzle about *for example*

One way to understand **imperatives** is as formally identifiable sentence types that are prototypically used for requesting or commanding. These prototypical functions as well as more peripheral usages of the same sentence type (e.g. wishing, advising) express a restriction of the possible course of events such that what is requested, commanded, advised or wished for is true. Therefore, they are all naturally linked to necessity in semantics, and it seems straightforward to interpret an imperative $\phi!$ as constraining all accessible future courses of events to $\phi$-courses (e.g. Asher and Lascarides 2003, Mastop 2005, Franke 2005). This makes straightforward predictions for most instances of imperatives and can even be extended to cover the somewhat marked permission usages as an indirect way of using necessity statements (cf. Schwager 2005b). Nevertheless, it fails to cover one reading of imperatives modified by *zum Beispiel* 'for example' in German, cf. (1).

(1) Kauf zum Beispiel keine Zigaretten!
buy.IMP for example no cigarettes
'For example, don't buy any cigarettes.'

Example (1) is ambiguous. As an answer to questions as in (2-a), it can be paraphrased as in (2-b). As an answer to (3-a), as in (3-b):

(2) a. $Q_1$: How could I stop smoking?/$Q_1$': What do I have to do in order to stop smoking?
b. One of the things you may not do is buy cigarettes. $\Box\neg BC(addressee)$
($\rightarrow$ *It is necessary that you don't buy cigarettes.*)

(3) a. How could I save money?
b. One of the things you could do is not buy cigarettes. $\Diamond\neg BC(addressee)$
($\not\rightarrow$ *It is necessary that you don't buy cigarettes.*)

So, (1) can either express that *not buying cigarettes* is part of the addressee's obligations, or that *not buying cigarettes* is a possibility to achieve the goal. On the second reading, not buying cigarettes is clearly not necessary. A semantics that relies on necessity fails to account for the reading exemplified in (3). The two variants of (2-a) show that the modal force is not automatically determined by the modal force of the question predicate ($Q_1$ contains possibility, $Q_1$' necessity as a question predicate; nevertheless, (1) is interpreted along the lines of (2-b) in both cases, that is, as expressing necessity).

The reading under which (1) is similar to (2-b) expresses that buying cigarettes is an **inexhaustive necessity** (that is, one obligation among others). The reading under which (1) is similar to (3-b) expresses that buying cigarettes is an **inexhaustive possibility** (that is, one possibility among others).

Before setting out for an analysis, it might be useful to take a look at their exhaustive counterparts. Example (4) displays **exhaustive possibility**:

(4) a. Q: What could I possibly do to stop smoking?
b. A: Du kannst nur aufhören, Zigaretten zu kaufen.
you can only stop, cigarettes to buy
'The only thing you can do is stop buying cigarettes.'

Example (4-b) expresses that the only possibility for the addressee to stop smoking is not to buy cigarettes anymore. The overt exhaustifier *only* is used to indicate exhaustivity. Consequently, if she wants to stop smoking, it is necessary that she doesn't buy cigarettes anymore. So, exhaustive possibilities come out as necessities that are not specified with respect to their degree of exhaustivity.

The unmodified necessity modal in (5) allows for an interpretation as **exhaustive necessity**. That is, given the task of getting into a good university, nothing is necessary apart from having a lot of money. The possibility of B's incredulous question clearly confirms the existence of such an interpretation.[1] But when overt *for example* forces a reading of inexhaustive necessity, B's incredulous question is completely incoherent (A's utterance has already indicated that having a lot of money may not be the only requirement to get into a good university).

(5) a. A: To get into a good university, you must have a lot of money. B: Really? And that's all?

[1] Nevertheless, it is most likely not part of the asserted proposition, as shown by B's correction in (i-a). Making exhaustive necessity explicit is not so easy though. Adding the exhaustifier *only* results in the **sufficiency modal construction** (cf. von Fintel and Iatridou 2005), cf. (i-b). But this does not only express that there are no other requirements than having enough money, but also that having enough money is ranked low on the scale of efforts.

(i) a. A: To get into a good university, you must have a lot of money. B: Yes, but there is more to it than that!
b. To get into a good university, you only have to have lots of money.

b. A: To get into a good university, you must for example have a lot of money. B: #Really? And that's all?

## 2. The proposal: diamonds for imperatives

In order to explain the ambiguity in (1), I want to argue that semantically imperatives express possibility with respect to a contextually given set of possible worlds. For the moment, I abstract away from their inherently non truth conditional character and treat them as modalized propositions.[2]

Possibility and necessity (as expressed also by modals like *must* and *may*) are analysed as propositional quantifiers relating a background and a complement propositions (cf. Geurts 1999). The modal element in an imperative $\phi!$ is assumed to consist in an imperative operator $OP_{Imp}$ (cf. (6-c)). Its background $b$ is typically interpreted as referring to the set of those worlds in the Common Ground that comply best with what the speaker wants, or in which the addressee reaches his current goal in a convenient way.

(6) a. $\Diamond = \lambda b \lambda p.(\exists w \in b)[w \in p]$
b. $\Box = \lambda b \lambda p.(\forall w \in b)[w \in p]$
c. $OP_{Imp} = \Diamond$

Exhaustivity and antiexhaustivity can now be treated as modifiers on propositional quantifiers. Both are of type $<<st,<st,t>>,<st,<st,t>>$ ($s$ and $t$ for worlds and truth values respectively).

*Being an exhaustive possibility with respect to background b*, $(EXH(\diamond))(b)$, can now be interpreted as *covering all of b*. This follows Zimmermann 2000's closure condition on lists of possibilities (cf. (7))[3]. Added to a list of possibilities $p_1, \ldots, p_n$, (7) expresses that this list is exhaustive in that the entire background $b$ is covered by their union. (8) simplifies it to an operator over single possibilities, which (for non-empty backgrounds) gives us the equivalence in (9).[4]

(7) $(\forall q)[q \cap b \neq \emptyset \rightarrow [q \cap p_1 \neq \emptyset \vee \ldots \vee q \cap p_n \neq \emptyset]]$ his ($24\kappa'$)

(8) $EXH(\Diamond) = \lambda b \lambda p.\Diamond(b)(p)$ & $(\forall q \in \Diamond(b))[q \in \Diamond(p)]$

(9) $EXH(\Diamond)(= EXH(OP_{Imp})) = \Box$

[2] Cf. Schwager 2005a for an elaboration of an additional presuppositional meaning component of imperatives that explains the inaccessibility of truth values.

[3] Zimmermann 2000 argues that for domains with mereological structure of propositions or locations, a simple general exhaustivity operator as proposed e.g. in Groenendijk and Stokhof 1984 cannot be applied. Although I cannot elaborate on this here, a more complex variant that takes into account comparative relevance (e.g. in terms of utility, cf. van Rooij and Schulz ta) should in principle be extendable to exhaustification with respect to properties like *being permitted* as well.

[4] For arbitrary $b(\neq \emptyset)$ and $p$: (*EXH*$(\Diamond)$)$(b)(p) \Leftrightarrow \Box(b)(p)$. Proof: $\Rightarrow$ If $w \in b$, then $\{w\} \cap b \neq \emptyset$; but then, $\{w\} \cap p \neq \emptyset \rightarrow w \in p$. $\Leftarrow$ For non-empty $b$, $\Diamond(b)(p)$ follows. And if $w \in q \cap b$, then $w \in p$. Hence, $q \cap p \neq \emptyset$. (This is an adaptation of Zimmermann's proof for lists of possibilities.)

Now, we have to generalize the notion of exhaustivity of a modal relation from possibility to covering also necessity. *p is an exhaustive necessity with respect to background b*, (*EXH*($\Box$))($b$)($p$), shall be interpreted as *nothing follows from the background b that doesn't follow from p*.[5]

(10) $\textit{EXH}(\Box) = \lambda b \lambda p.\Box(b)(p)$ & $(\forall q \in \Box(b))[q \in \Box(p)]$

From (8) and (10), we can generalize to the following modifier *EXH* of propositional quantifiers $R$:

(11) $\textit{EXH}(R) = \lambda b \lambda p.R(b)(p)$ & $(\forall q \in R(b))[q \in R(p)]$

A natural interpretation for the antiexhaustifier *zum Beispiel* 'for example' is to assume that it modifies a propositional quantifier by adding that the speaker doesn't exclude that other propositions than the expressed argument proposition stand in the same relation to the background. This is spelled out in (12).

(12) $\textit{zB}(R) = \lambda b \lambda p.R(b)(p)$ & $\Diamond(Bel_S)[\neg(\textit{EXH}(R))(b)(p)]$,
where $\textit{Bel}_S$ is the set of the speaker's belief worlds.

So, for instance, if $p \in$ (*zB*($\Box$))($\cap$*what is commanded*), then $p$ is an obligation, but the speaker doesn't exclude that there are further obligations independent of $p$.

Semantically, the imperative operator $OP_{Imp}$ is equivalent to the modal verb *may*. Nevertheless, it differs in its interaction with (anti)exhaustification. $OP_{Imp}$ combines obligatorily either with overt *zB* or with covert *EXH* (default). Only after doing so, it behaves like a modal in optionally combining with *EXH* or *zB*, before applying to background and complement proposition. The possible LF-schemata are given in (13) ($\emptyset$ indicates the absence of an (anti)exhaustifier at the respective position, and options are indicated in curly braces).

(13) a. [ [ {*EXH*, *zB*, $\emptyset$} [ {*EXH*, *zB*}($OP_{Imp}$) ] ] $b$ $p$ ]
b. [ [ {*EXH*, *zB*, $\emptyset$} [ {*must,may,*... } ] ] $b$ $p$ ]

According to (13-a), in the absence of *zum Beispiel*, *EXH* has to apply to $OP_{Imp}$. Consequently, by the equivalence in (9), possibility is turned into necessity, giving the desired necessity reading for plain imperatives.

The ambiguity of (1) relies on the two positions available for *zB* with respect to $OP_{Imp}$ (cf. (13)). If *zum Beispiel* serves as the obligatory modifier of $OP_{Imp}$, the imperative expresses true possibility, cf. (14).

(14) $\textit{zB}(OP_{Imp}) = \lambda b \lambda p.\Diamond(b)(p)$ & $\Diamond(\textit{Bel}_S)\neg(\forall q \in \Diamond(b))[q \in \Diamond(p)]$

According to (14), (*zB*($OP_{Imp}$))($b$)($p$) says that $p$ is a possibility with respect to background $b$, but that the speaker holds it possible that parts of $b$ are not covered by

[5] Most likely logical consquence is too strong and should ultimately be replaced by a context sensitive consequence relation.

$p$ (consequently, that $p$ is not a necessity). This accounts for reading (3-b).

The computation for the inexhaustive necessity reading individuated in (2-b) is a bit more complicated. In (16), *EXH* has applied to $OP_{Imp}$ and has turned it into necessity, while *zB* occupies the position of the optional modifier above. Under the common pragmatic assumption spelled out in (15), this accounts for the reading of inexhaustive necessity.

(15) For any speaker S and any proposition A: $\text{utter}_S(A) \rightarrow \Box(\textit{Bel}_S)A$.

(16) *zB*(*EXH*($OP_{Imp}$)) = by *zB* in (12)

a. $\lambda b\lambda p.(\textit{EXH}(\Diamond))(b)(p)$ & $\Diamond(\textit{Bel}_S)[\neg(\textit{EXH}(\textit{EXH}(\Diamond)))(b)(p)] =$ by (9), (10)

b. $\lambda b\lambda p.\Box(b)(p)$ & $\Diamond(\textit{Bel}_S)[\neg[\Box(b)(p)$ & $(\forall q \in \Box(b))[q \in \Box(p)]]$

By (15) and the first conjunct of (16-a), we know that $\neg\Diamond(\textit{Bel}_S)\neg\Box(b)(p)$. By De Morgan's law, the last conjunct in (16-b) can then be simplified so as to give us (17):

(17) $\lambda b\lambda p.\Box(b)(p)$ & $\Diamond(\textit{Bel}_S)[\neg(\forall q \in \Box(b))[q \in \Box(p)]]$

So, *zB*(*EXH*($OP_{Imp}$))$(b)(p)$ says that $p$ is a necessity with respect to $b$, but that the speaker does not exclude that further, independent propositions are $b$-necessities as well. This is exactly the reading of inexhaustive necessity we are after for (2-b).

Moreover, it is predicted correctly that application of *EXH* to any $R$ that has been antiexhaustified by *zB* results in attributing contradictory beliefs to the speaker.

(18) $^{\#}$*EXH*(*zB*($R$)) = by (11)

a. $\lambda b\lambda p.(\textit{zB}(R))(b)(p)$ & $(\forall q \in (\textit{zB}(R)(b)))[q \in (\textit{zB}(R))(p)] =$ by (12)

b. $\lambda b\lambda p.R(b)(p)$ & $\Diamond(\textit{Bel}_S)[\neg(\forall q \in R(b))[q \in R(p)]]$ &
$(\forall q \in \{t \mid R(b)(t)$ & $\Diamond(\textit{Bel}_S)[\neg(\forall q' \in R(b))[q' \in R(t)]]\})$
$[q \in \{s \mid R(p)(s)\&\Diamond(\textit{Bel}_S)[\neg(\forall q' \in R(p))[q' \in R(s)]]\}]$

For arbitrary $b$ and $p$, the last conjunct causes the contradiction. Insert $p$ as a $q$. Due to the first two conjuncts, $p$ passes the restriction $(R(b)(p)$, and $\Diamond(\textit{Bel}_S)[\neg(\forall q' \in R(b))[q' \in R(p)]])$. Consequently, it should hold that $R(p)(p)$ - which might be true or not, depending on the nature of $R$, but crucially that $\Diamond(\textit{Bel}_S)[\neg(\forall q' \in R(p))[q' \in R(p)]]$. Hence, applying *EXH* to an operator that has been antiexhaustified by *zB* attributes nonsensical beliefs to the speaker and is therefore most likely avoided.

## 3. Conclusion and outlook

*EXH* and *zB* as defined here allow us to compute the different modal forces observed with imperatives depending on the interaction of $OP_{Imp}$ with *zum Beispiel*. This can't be obtained if imperatives are interpreted as always expressing necessity. *EXH* and *zB* carry over to modal verbs as well.

So far, this all happens in semantics, which is maybe not as it should be, especially if we take serious the observations concerning modal verbs. Further unification with other approaches to exhaustification and work on *only* remains to be done.

Empirically, it would be interesting to compare the proposal with exhaustivity in disjunctions (cf. Geurts ta), and to try to extend it to modal operators in Salish that (like imperatives) express necessity as a default but are interpreted as possibility when necessity gives rise to a contradiction (cf. Matthewson et al. 2005). Last but not least, the assumption of an exhaustivity operator in the imperative might shed new light on the interaction of imperatives with free choice items (cf. Menèndez-Benito 2005 for licensing of free choice items in connection with exhaustification).

**Acknowledgements**

For discussion and helpful comments I want to thank Rick Nouwen and Ede Zimmermann. Of course, the usual disclaimer applies.

**Bibliography**

Asher, N. and Lascarides, A.: 2003, Imperatives in dialogue, in P. Kuehnlein, H. Rieser, and H. Zeevat (eds.), *Perspectives on Dialogue in the New Millennium*, Benjamins, Amsterdam

von Fintel, K. and Iatridou, S.: 2005, *Anatomy of a Modal*, Ms., MIT

Franke, M.: 2005, *Pseudo-Imperatives*, MAthesis, ILLC Amsterdam.

Geurts, B.: 1999, *Presuppositions and Pronouns*, Elsevier, Oxford

Geurts, B.: t.a., *Entertaining Alternatives*, To appear in 'Natural Language Semantics'

Groenendijk, J. and Stokhof, M.: 1984, *Studies on the Semantics of Questions and the Pragmatics of Answers*, Ph.D. thesis, Amsterdam

Mastop, R.: 2005, *What can you do?*, Ph.D. thesis, ILLC Amsterdam

Matthewson, L., Rullman, H., and Davis, H.: 2005, *Salish Modal Operators*, Talk at ESSLLI 2005, Edinburgh

Menèndez-Benito, P.: 2005, *Freedom of Choice*, Talk at 'Sinn und Bedeutung 10', 13.-15. 10. 2005, Berlin

Schwager, M.: 2005a, *Interpreting Imperatives*, Ph.D. thesis, University of Frankfurt/Main

Schwager, M.: 2005b, Permitting permissions, in J. Gervain (ed.), *Proceedings of the 10th ESSLLI Student Session 2005, Edinburgh*

van Rooij, R. and Schulz, K.: t.a., *Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation*, To appear in 'Linguistics and Philosphy'

Zimmermann, T. E.: 2000, Free choice disjunctions and epistemic possibility, *Natural Language Semantics* 8, 255–290

# WORD MEANING, UNIFICATION AND SENTENCE-INTERNAL PRAGMATICS

TORGRIM SOLSTAD

University of Oslo
torgrim.solstad@ilos.uio.no

Recent important developments within Discourse Representation Theory include a more elaborate formalisation and account of presuppositional phenomena, as well as the integration into the theory of unification as a mode of composition. Focusing on these issues, the following claims are made: (i) the varying compositional impact of some adverbials, ranging from merely constraining the properties of a predicate to radically altering them, is suitably modeled applying unification, and (ii) pragmatic mechanisms like bridging, presupposition verification and accommodation can apply mainly sentence-internally for some lexical items. To substantiate these claims, the analysis will centre around the German causal preposition 'durch' ('through').

## 1. Introduction

An adverbial can be said to be a free syntactic constituent which modifies a predicate semantically. However, some adverbials not only modify a predicate, but may even (radically) alter its properties. Prepositional adjuncts headed by the German causal-instrumental preposition *durch* ('by', 'through') are examples of one such type of adverbial. One of the main functions of *durch* is marking its complement as the causing event in a causal relation between two events, as exemplified in (1) and (2):

(1) *Der Polizist wurde getötet durch einen Schuss aus der eigenen Dienstwaffe.*
'The policeman was killed by a shot from his own service weapon.'

(2) *Der Polizist starb durch einen gezielten Schuss.*
'The policeman died through an accurate shot.'

In (1), the causative predicate *töten* ('kill') is used, which implies the existence of a causing event without specifying it, i.e. it is non-manner-specific. The modifying *durch* adjunct provides this specification: the death of the policeman is caused by the event of a shot from his own service weapon. In (2), the inchoative predicate *sterben* ('die') is used. Inchoatives like *sterben* are not generally assumed to imply a causative relation. Still, in combination with the *durch* adjunct, a semantics parallel to the one indicated for (1) is desirable: a shooting event is the cause of the policeman's death. Additionally, an inchoative like *sterben* does not associate with

an agent on its own. But sentence (2) clearly implies the presence of an agent, as the specification of the shooting event as being 'accurate', indicates. Thus, the *durch* phrase can be said to have altered the properties of the inchoative predicate *sterben*.

Accordingly, the semantics of both (1) and (2) can be represented as indicated in (3), leaving out information not relevant to the discussion here:

(3) $\lambda e_1 \exists e_2 [\text{BECOME}(\textit{dead}(p))(e_2) \wedge \text{CAUSE}(e_2)(e_1) \wedge \text{SHOOT}(e_1)]$

However, since inchoatives are not assumed to imply causation, there must be two different sources for the abstract predicate CAUSE: with causatives it originates in the predicate itself, but with inchoatives, the preposition seems to be the most plausible candidate for its introduction. But if *durch* in some cases should include a CAUSE of its own, principles of strict compositionality would seem to force us to assume an ambiguity between two *durch* prepositions since no two CAUSE predicates are assumed after the composition of *durch* with causatives. Assuming ambiguity would however clearly be somewhat counter-intuitive, given the parallel interpretation of (1) and (2). Thus, other means of composition for *durch* phrases and the predicates they modify, should be explored.

## 2. A unificational analysis

To deal with this challenge, the semantics of *durch* will be analysed by means of unification in Discourse Representation Theory (Bende-Farkas and Kamp 2001), applying principles of the presuppositional analysis of Kamp (2001, pp. 221-231) and Sæbø's (2005) analysis of *by*. Building on work by van der Sandt (1992), Kamp assumes that semantic information in a sentence is processed bottom-up via a storage algorithm. Semantic information represented preliminarily in the store part enters a main content part as it is bound, verified or accommodated. The general representational format of Kamp (2001) for a semantic node in a tree structure is shown in (4):

(4) $\left\langle \overbrace{\left\{ \langle \text{Variable}, \boxed{\text{Constraint}}, \text{Binding condition} \rangle \right\}}^{\text{STORE}}, \boxed{\text{CONTENT}} \right\rangle$

A semantic node representation consists of a pair of a content and a store element. The content is always a Discourse Representation Structure (DRS). The store is a set of one or more elements, each being a triple of a variable, a constraint (a DRS) and a binding condition. Binding conditions determine which variables can enter a binding relation, and constraints contain semantic information which may also be of importance for binding. In addition to the binding mechanism, a principle which unifies variables and constraints when possible, is assumed.

This machinery allows a unified analysis of the above uses of *durch* where the preposition indeed includes a CAUSE of its own. When combined with a causative

predicate, the implicit CAUSE of *durch* is not added to the content part since there is a CAUSE present in the predicate. However, the combination of *durch* with an inchoative leads to the projection of the CAUSE element in the content part. The actual formalisation is illustrated briefly below. *Durch* may be represented as in (5):

(5) **durch:**

$$\left\langle \left\{ \begin{array}{l} \langle e_1\ , \boxed{\begin{array}{l}\text{CAUSE}(e_2)(e_1)\\ e_1 \subseteq t_{loc}\end{array}}\ , \lambda_1\rangle, \\ \langle e_2, \boxed{\text{CAUSE}(e_2)(e_1)}\ , \lambda_2\rangle, \\ \langle t_{loc}, \qquad\qquad , \text{loc.t.}\rangle \end{array} \right\}, \begin{array}{|c|}\hline \\ \hline \\ \\ \hline \end{array} \right\rangle$$

*Durch* has no content of its own – its content part is empty –, but includes two event variables and a temporal variable in the store. The two event variables are further specified as entering a CAUSE relation. The binding conditions $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ indicate that the variables need to bind. When the complement of the preposition is added, as in *durch* ***einen Schuss***, the event expressed therein is bound by $e_1$ and the information in the noun is added as a further constraint on the causing event: $\text{SHOT}(e_1)$ (cf. Chung & Ladusaw (2004), where the term *restriction* is used). When a *durch* phrase is combined with a causative predicate which has a completely parallel store part, the variables of *durch* and their constraints will eventually be unified with or bind the variables of the causative predicate. The causing event $e_1$ of *durch*, which has already bound the event in the complement of *durch*, will be unified with the causing event of the predicate, whereas the caused event $e_2$ will bind the caused event of the predicate. Additionally, the constraints of the predicate and the preposition are merged and - where applicable - unified. After binding and unification have occured, the actual contribution of a *durch* phrase, as compared to the information provided by the predicate alone, is restricted to the specification of the causing event given by the constraint $\text{SHOT}(e_1)$. Turning next to the inchoative predicate, its store part includes only one event, which will be bound by $e_2$ of *durch*. In this case, the variable of the causing event of *durch* will be added to the content, since there is no event for it to be unified with. Furthermore, the CAUSE relation of which the bound event variable of the inchoative predicate will be a part, will also enter the content, along with the aforementioned constraint derived from the complement of the preposition.

## 3. Sentence-internal pragmatics

This treatment of *durch* amounts to analysing its implicit CAUSE element as an *intrasentential* presupposition. A *durch* phrase can be said to *assert* the event included therein and *presuppose* that this event is a cause of some other event. The common basis for generally assumed mechanisms for presuppositional behaviour and the compositional unification-based analysis of *durch* is as follows: When combined with causatives, *durch* seems to lack a meaning of its own. This is due to the unifi-

cation of the CAUSE of *durch* with the CAUSE of the predicate, which is parallel to presupposition verification. In combination with inchoatives, however, *durch* does seem to make a greater contribution, where a CAUSE predicate is introduced by the causal preposition itself. Here, a parallel to context accommodation can be observed.

Importantly, a pragmatic account of the combinatorial potential of *durch* can capture some further properties of the preposition which have previously been ignored or not correctly identified. Two additional pragmatic mechanisms involved are *bridging* and *acceptability*. In (6), bridging (in the wider sense of Bittner (2001)) can be argued to take place, where the CAUSE associated with the preposition forces a reinterpretation of the state described in the predicate *hoch* ('high') as being a caused resultant state:

(6) *Auch der durch diese Haltung hohe Luftwiderstand kann auf längeren Strecken ganz schön schlauchen.*
'The high air resistance due to this posture may put you through the mill over longer distances.'

In (7), it can be seen that claims made in the literature that *durch* generally cannot be combined with manner-specific causatives (Härtl 2001) are not correct:

(7) *Er wurde* $\left\{ \begin{array}{l} {}^{??}\textit{durch einen Schuss} \\ {}^{\text{OK}}\textit{durch einen Genickschuss} \end{array} \right\}$ *erschossen.*
'He was shot dead $\left\{ \begin{array}{l} {}^{??}\text{with a shot} \\ {}^{\text{OK}}\text{with a shot to the neck} \end{array} \right\}$.'

The well-formedness of such combinations should not be explained by reference to the semantics of *durch*. A more general account of the distribution in (7) is achieved by assuming that composition is restrained by a general pragmatic mechanism of acceptability as described by van der Sandt (1992, pp. 367 ff.). Modifying a predicate such as *erschießen* ('shoot dead') by an adjunct like *durch einen Schuss* ('with a shot') is uninformative and thus unacceptable. However, a specification such as *durch einen* ***Genick****schuss* ('with a shot *to the neck*') renders the adjunct more specific than the shooting event described in the predicate, adding to the content. Thus, the distribution of *durch* phrases in combination with manner-specific causatives does not bear on the semantics of *durch*, but is determined by acceptability restrictions.

It should be emphasised that in the examples above, all pragmatic mechanisms assumed to account for the compositional behaviour of *durch* apply purely sentence-internally. Since presuppositions in general are assumed to be verified also inter-sententially, *durch* might seem like an exception. But there is at least one type of occurence where the presupposition of *durch* can be seen as being verified sentence-externally:

(8) *Sie hat Geld verloren. Es geschah durch Unaufmerksamkeit.*
'Sie lost some money. It happened due to lack of attention.'

In the second sentence in (8) containing the *durch* phrase, the abstract event predicate *geschehen* ('happen') is used, which asserts that some event took place. What *durch* modifies semantically however, is the predicate *verlieren* ('lose') in the first sentence. Thus, in the case of (8), part of the presuppositional information in the store of *durch* binds an event variable in the preceding sentence.

## 4. A wider perspective

An approach as sketched above has applications beyond the analysis of *durch*. First, unification as a mode of composition has been applied in an analysis of the semantics of *by* in English (Sæbø 2005). Second, there are causal prepositions in other languages which show a similar behaviour to *durch*. In English, *through* can also be combined with both causative and inchoative predicates. More interestingly, given the close relationship between English *through* and German *durch*, a language more remotely related to German such as Bulgarian also has a preposition, *ot* ('from'), which combines with causatives and inchoatives:

(9) a. *Toj be ubit ot tri kurshuma.*
He was killed from three shots
'He was killed with three shots.'

b. *Toj sagina ot tri kurshuma.*
He died from three shots
'He died from three shots.'

Third, there are other types of adverbial modification, where the above analysis can be applied plausibly, as illustrated in (10):[1]

(10) a. *Sie ging in das Haus hinein.*
DIR+IN DIR+IN
'She went into the house.'

b. *Sie ging in das Haus.*
'She went into the house.'

c. *Sie ging hinein.*
'She went inside.'

In (10a) the adverbials *in das Haus* ('into the house') and *hinein* ('inside' in addition to viewpoint information) specify a single path of movement. They are not interpreted as describing two paths which are combined. There is a double specification of an *in* movement (i.e. *into* as opposed to *out of*), once in the preposition *in* and once in the *hinein* element. In addition, directionality is specified twice: once in the combination of the preposition with accusative case, and once in the *hinein* element.

[1] Thanks are due to Christopher Habel for pointing my attention to this example.

As can be seen from (10b-c), either of the advberbials in (10a) can occur without the other. In the spirit of the analysis presented here, the *hinein* element would be assumed to carry the presupposition that there is an object into which movement takes place. In (10a) this presupposition is sentence-internally verified, whereas it will have to be verified in a wider context or accommodated in (10c). The information on directionality and inwards movement of the two adverbials is unified whenever they both occur.

In sum, these data suggest that the presuppositional analyses of Kamp (2001) and van der Sandt (1992) in combination with unification-based composition can be suitably applied in analysing lexical items other than e.g. particles and factive verbs.

## 5. Conclusion

In this paper, it was argued that unification is an adequate mode of composition in accounting for the varying compositional impact of adverbials. It was also argued that pragmatic mechanisms are important in describing the combinatorial distribution of some lexical items.

**Acknowledgements**

I would like to thank Cathrine Fabricius-Hansen, Hans Kamp, Elena Karagjosova, Manfred Krifka, Kjell Johan Sæbø and Henk Zeevat for valuable comments.

## Bibliography

Bende-Farkas, A. and Kamp, H.: 2001, *Indefinites and Binding: From Specificity to Incorporation*, Revised version of the lecture notes from a course given at the 13th ESSLLI in Helsinki, Finland

Bittner, M.: 2001, Surface composition as bridging, *Journal of Semantics* 18(2), 127–177

Chung, S. and Ladusaw, W. A.: 2004, *Restriction and Saturation*, MIT Press, Cambridge, Massachusetts

Härtl, H.: 2001, CAUSE *und* CHANGE*: thematische Relationen und Ereignisstrukturen in Konzeptualisierung und Grammatikalisierung*, Akademie-Verlag, Berlin

Kamp, H.: 2001, The Importance of Presupposition, in C. Rohrer, A. Roßdeutscher, and H. Kamp (eds.), *Linguistic Form and its Computation*, pp 207–254, CSLI Publications, Stanford

van der Sandt, R. A.: 1992, Presupposition Projection as Anaphora Resolution, *Journal of Semantics* 9(4), 333–377

Sæbø, K. J.: 2005, The structure of criterion predicates, in J. Dölling and T. Heyde-Zybatow (eds.), *Event Structures in Linguistic Form and Interpretation*, Mouton de Gruyter, Berlin

# CAUSATIVE CONSTRUCTIONS AND ASPECTUAL MEANINGS: A CASE STUDY FROM SEMITIC DERIVATIONAL MORPHOLOGY

REUT TSARFATY

Institute for Logic, Language and Computation
University of Amsterdam
rtsarfat@science.uva.nl

This work aims at identifying aspectual properties of events denoted by morphological causatives in Modern Hebrew (MH). The main purpose of this investigation is to establish a clear connection between causative constructions and aspectual meanings, two notions that are not so easily correlated. A secondary goal is to argue for the systematic aspectual contribution of Semitic derivational morphology. Our theory is inspired by Smith's causal chain and builds on a thematic account of Semitic derivational morphology. Combining a formal and empirical investigation we argue that the MH causative template *Hiph'il* shifts the viewpoint of an event onto its initiation and development phases, making it more appropriate for imperfective use.

## 1. Causative Constructions and Aspectual Meanings

MH exhibits three kinds of causative expressions that can be found across languages: (i) lexical causatives (ii) morphological causatives, and (iii) periphrastic causatives, illustrated in (1). The latter two kinds are termed causative constructions. Morphological causatives in MH are derived by fusing consonantal roots with the consonant/vowel skeleton known as the causative template *Hiph'il*.

| | | | |
|---|---|---|---|
| (1) | Fred harag et Bill | Fred hemit et Bill | Fred garam leBil lamut |
| | Fred killed ACC Bill | Fred die.*Hiph'il* ACC Bill | Fred caused to-Bill to-die |
| | Fred killed Bill | Fred made Bill dead | Fred caused Bill to die |

Morphological causatives in MH give rise to a wide range of meanings that does not necessarily coincide with causation in its strict sense. For example, the MH causative verb $he'e\check{k}il$ (feed) is derived from the root $[a][\check{k}][l]$ (eat) although it is debatable whether 'feed' is equivalent to 'cause to eat'. This work aims at identifying the aspectual properties shared by such derived verbs.

As of yet, research into lexical semantics has not shown a systematic correlation between causative constructions and aspectual meanings. In particular, Levin 2000 shows that causatives cannot be reduced to any one kind of *Aktionsart*. However, causatives are valence increasing operations thereby encoding speakers' choice to incorporate an additional element (a cause) into the event description. This makes

them aspectually marked as well. Further, causatives are typically analyzed as complex situations consisting of two, causally related, events. Yet, Song 2001 makes the typological observation that the causing event is highly abstracted and has no specific lexical content. This allows us to restrict our analysis to the caused predicate and contrast its aspectual meaning with the respective unmarked predicate.

The key idea is that internal structure of events and the thematic content of their participants are intertwined. We view participants as elements in the temporal extension of an event, and show how marked thematic relations affect aspectual content. Such an approach allows one to formally investigate aspectual meanings in languages that do not grammaticalize aspect, yet mark thematic relations by formal means.

## 2. The Causal Chain

To make the desired link between aspectual meanings and causative constructions we devise a proposal inspired by the *causal chain* presented in Smith 1991. Smith uses the causal chain to distinguish aspectual classes (*Aktionsarten*) of verbal expressions based on how much of the chain is covered. The order from left to right is iconic to temporal precedence: CAUSE SUBJECT ACTION OBJECT RESULT[1]. Smith's schematic description is informal and does not allow to make precise predictions. In particular, it is unclear which situations are characterized by the CAUSE element.

To remedy this, I first draw a distinction between a 'cause' relation and a 'CAUSE' element. A 'cause' relation is a relation between any two adjacent elements in the causal chain. In the current account, a 'cause' relation entails temporal precedence. Using a simplified chain and the 'cause' relation we can represent the Aktionsarten.

(2) a. State ..... RESULT
b. Activity .....SUBJECT ACTION.....
c. Achievement ..... ACTION OBJECT RESULT
d. Accomplishment SUBJECT ACTION OBJECT RESULT

Causative constructions contribute a 'CAUSE' element which is crucially disjoint from elements already existing in the representation of a given situation[2]. The 'CAUSE' stands in 'precedes' and 'cause' relations to the elements in the given chain. Since not all situations map onto the entire span, the emergent chains give rise to the event interpretations in (3). The 'CAUSE', which is lexically underspecified, extends the representation to include the immediately preceding element.

(3) a. A caused state .....CAUSE RESULT
b. A caused activity CAUSE SUBJECT ACTION .....

[1] We collapse INSTRUMENT into ACTION as they temporally overlap (cf. Smith 1991, page 34).
[2] I use the term CAUSE in its broad sense here, and the participant associated with the 'CAUSE' need not be an instigator and not even volitional. The observation which is pertinent to the aspectual account is that the 'CAUSE' element serves as a precondition for the initiation/progression of the caused event.

c. A caused achiev. .....CAUSE ACTION OBJECT RESULT

d. A caused accomp CAUSE SUBJECT ACTION OBJECT RESULT

Further, marking a 'CAUSE' focuses the linguistic description on the forces behind the initiation and development of the event, thus altering its aspectual viewpoint. A similar shift characterizes periphrastic constructions using the aspectual verbs 'start' and 'continue'. Smith 1991 terms such morphemes *super-lexical morphemes* as they "modulate the focus of a situation rather then determining the situation itself".

## 3. Modern Hebrew Morphological Causatives

Verbs in MH are derived from tri-consonantal roots plugged into templates of consonant/vowel skeletons termed *binyanim*. Doron 2003 argues that the templates alter thematic relations in a predictable way. Following Dowty 1991, it is widely accepted that thematic properties are selected in accord with the event denoted by the verb. Here we hypothesize that the converse also holds, i.e. that the event structure of a derived verb changes to accord with the altered thematic relations.

Similar to Doron 2003, we assume a narrow lexicon consisting of coarse-grained roots. Roots have basic meanings that can be *approximated* by plugging them into the so-called 'simple template' (a.k.a. *Pa'al*) which is morphologically and thematically unmarked. These basic meanings induce a preliminary classification of Aktionsarten. Doron shows that the causative template contributes an external participant that serves as the cause to the event at hand. The addition of an external participant that stands in a 'cause' and 'precedes' relation alters the event structure as described in the previous section, thus shifting its aspectual meaning in a predictable way.

## 4. Formal Account

We formalize the theory using the Event Calculus (EC) of van Lambalgen and Hamm 2005, a formalism to reason about time and change that axiomatizes cause/effect relations. EC requires (at least) events ($e, e'..$), time instants ($t_1, t_2..$), and time dependent properties called *fluents* ($f_1, f_2..$). The schematic representation of aspectual classes (*eventualities*) in EC bears striking resemblance to our revised chain.

(4) An *eventuality* is a structure $\langle f_1, f_2, e, f_3 \rangle$ where

a. $f_1$ represents an activity which exerts a force,

b. $f_2$ represents a changing object/state driven by the force of $f_1$,

c. $e$ represents a canonical goal, and

d. $f_3$ represents the state of having achieved the goal.

*Aktionsarten* are defined in EC using these quadruples ( $[.]$ indicates the viewpoint)[3]

[3]This formal representation is over-simplified. The eventuality quadruple is, in fact, an abbreviation for a *scenario*; a sequence of general statements universally quantified with respect to time that, together

(5) a. *States* (e.g. love, know) $\langle -, -, -, [f_3]\rangle$
b. *Activities (wide)* (e.g. walk, push) $\langle [f_1], f_2, -, -\rangle$
c. *Achievement* (e.g. fall, break) $\langle -, -, [e], f_3\rangle$
d. *Accomplishments* (e.g. build, create) $\langle [f_1], f_2, e, f_3\rangle$

### 4.1. The Causative Template *Hiphil*

Fusing a root with the causative template has the effect of filling in preceding empty slots, and shifting the viewpoint to the causing element. This gives rise to altered event representations. This proposal accounts for the aspectual meanings of a wide range of morphologically derived causatives in MH, as illustrated in (6)[4].

(6) a. *State* $\langle -, -, -, [f_3]\rangle \rightsquigarrow$ *Inchoative state* $\langle -, -, [e], f_3\rangle$
1. $[d][a][g]$ + ***Pa'al*** $= da'ag$ (be worried)
2. $[d][a][g]$ + ***Hiph'il*** $= hid'id$ (make worry)
b. *Achievement* $\langle -, -, [e], f_3\rangle \rightsquigarrow$ *Progressive achievement* $\langle [f_1], f_2, e, f_3\rangle$
1. $[n][\check{p}][l]$ + ***Pa'al*** $= na\check{p}al$ (fall)
2. $[n][\check{p}][l]$ + ***Hiph'il*** $= hepil$ (fell, made fall, cause to fall)
c. *Activity* $\langle [f_1], f_2, -, -\rangle \rightsquigarrow$ *Ingressive activity* $\langle [f_{0_1}], f_{0_2}, [f_1], f_2, -, -\rangle$
1. $[r][k][d]$ + ***Pa'al*** $= rakad$ (dance)
2. $[r][k][d]$ + ***Hiph'il*** $= hirkid$ (cause to dance, made dance)
d. *Accomplish.* $\langle [f_1], f_2, e, f_3\rangle \rightsquigarrow$ *Ingressive accomplish.* $\langle [f_{0_1}], f_{0_2}, [f_1], f_2, e, f_3\rangle$
1. $[a][\check{k}][l]$ + ***Pa'al*** $= a\check{k}al$ (eat)
2. $[a][\check{k}][l]$ + ***Hiph'il*** $= he'e\check{k}il$ (feed)

In (6a), the state of 'being worried' comes about due to a certain cause, which gives the event an *inchoative* interpretation. In (6b), the event 'fall' is extended to include a preparatory phase that precedes and causes it, giving it the interpretation of a *progressive achievement* (which mirrors an accomplishment). In (6c) and (6d), the durative events are extended to include a preceding and parallel cause that continuously stimulates the 'caused' event, which provides it with an *ingressive* interpretation.

The same proposal accounts for the aspectual meanings of *denominal* causatives, i.e. causative verbs that are derived from nouns (which crucially do not exhibit an event structure of their own). We identify the object denoted by the noun with $f_2$, the template fills in $f_1$ with the essential 'cause', and the result gives rise to a variety of wide activities, including the emission verbs mentioned in Doron 2003, e.g. (7).

with the EC axioms, defines the micro-theory of the event. Grammatical viewpoints are formalized using *integrity constraints* which relate the *reference point* of the eventuality to one of its components (see van Lambalgen and Hamm 2005). The default viewpoints presented here are specific to MH.

[4]Filling in a changing fluent $f_2$ always requires filling in its driving force $f_1$, in which case the 'CAUSE' element turns out to be a complex element $\langle f_1, f_2, -, -\rangle$, referred to in EC as a *dynamics*. The same complex element is required when the caused event is a durative one (i.e., activities) that already involve a dynamics. See Tsarfaty 2005 for the complete formalization.

(7) *noun* $\langle -, [f_2], -, - \rangle \rightsquigarrow$ *activity* $\langle [f_1], f_2, -, - \rangle$

a. 1. $[r][']['][\check{s}] + noun = ra'a\check{s}$ (noise)

2. $[r][']['][\check{s}] + Hiphil = her'i\check{s}$ (emit noise)

The analysis of denominal causatives serves to demonstrate the two core components of our theory. First, that the addition of a new participant adds also the aspectual context in which it operates, and second, that Semitic derivational morphology has an indispensable aspectual contribution.

## 5. Empirical Investigation

To support our theoretical findings we set out to find empirical evidence for aspectual choice in MH and for the development of verb forms' usage. We asked 22 native MH speakers (ages 3–30) to narrate a story based on a wordless picture book from two different viewpoints. Once while walking through the pictures ('Part I'), and once in retrospect, after the successful resolution of the plot ('Part II').

Figures 1– 4 summarize the results of our investigation. Figure 1 shows that the use of simple verb forms in the narratives decreases with age. Figure 3 shows a respective increase[5] in the use of causative verb forms. This joint distribution indicates that adult-like use of the morphological templates requires a longer acquisition phase than mastering the grammatical tenses[6] (already achieved by the age of 3, Berman and Slobin 1994). Figures 2- 4 show, for all age groups, a persistent increase in simple verb forms used in 'Part II' *relative* to 'Part I', along with a respective decrease in causative verb forms. This shows a preference for causative verbs to describe ongoing/incomplete events as they happen, and for simple verbs to describe complete/completed events and drive the story time-line forward.

We conclude that simple verb forms are morphologically unmarked, and semantically unmarked with respect to particular elements in the internal structure of an event. Thus, they refer to events in their entirety, which makes them appropriate for perfective use. Causative verbs, on the other hand, make explicit reference to elements in the internal structure of the event (i.e., its cause), and focus on its initiation and development phases. Therefore, their aspectual value is semantically marked, which makes them appropriate for describing imperfective situations.

**Acknowledgements**

I am grateful to Prof. Michiel van Lambalgen for supervising this work and Darrin Hindsill for fruitful discussions throughout. As of 2005 my work is supported by the Netherlands Organization for Scientific Research (NWO), grant number 017.001.271.

[5] With the exception of the peak at the age of 4, which may be understood as an *over-generalization* spike.
[6] Note that tense marking is obligatory while marked choices of the kind discussed here are voluntary.

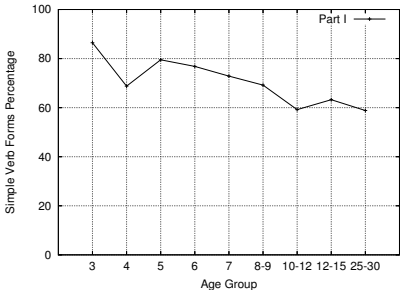



Figure 1: Percentage of simple verb forms used in the first story (average per age group)

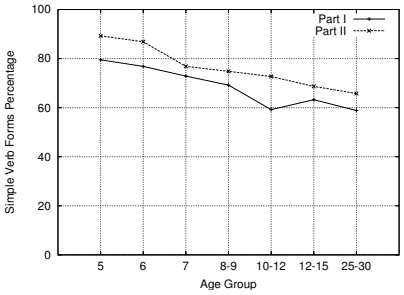



Figure 2: Percentage of simple verb forms used in each of the stories (avg. per age group)

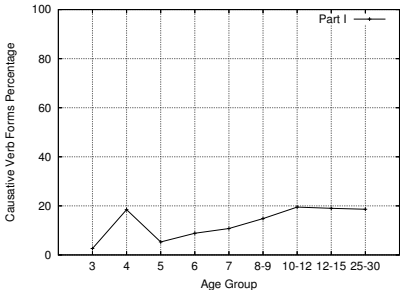



Figure 3: Percentage of causative verb forms used in the first story (average per age group)

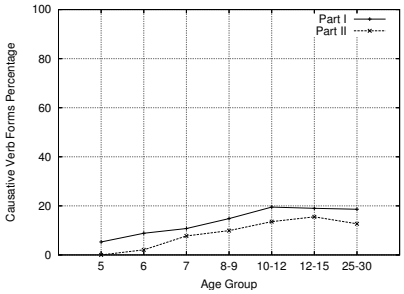



Figure 4: Percentage of causative verb forms used in each of the stories (avg. per age group)

## Bibliography

Berman, R. A. and Slobin, D. I.: 1994, *Relating Events in Narratives: A Crosslinguistic Developmental Study*, Lawrence Erlbaum, Hillsdale, New Jersey

Doron, E.: 2003, Agency and Voice: The Semantics of the Semitic Templates, *Natural Language Semantics* (11), 1–67

Dowty, D.: 1991, Thematic Proto-Roles and Argument Selection, *Language* 67(3), 547–619

Levin, B.: 2000, Aspect, lexical semantic representation, and argument expression, in *The 26th meeting of the Berkeley Linguistics Society, Proceedings*, pp 413–329

Smith, C. S.: 1991, *The Parameter of Aspect*, Kluwer, Dordrecht

Song, J. J.: 2001, *Linguistic Typology: Morphology and Syntax*, Chapt. 5, pp 257–296, Longman Linguistics Library, Pearson Education Limited, England

Tsarfaty, R.: 2005, 'binyanim ba'avir': An Investigation of Aspect Semantics in Modern Hebrew, *Master's thesis*, ILLC, University of Amsterdam

van Lambalgen, M. and Hamm, F.: 2005, *The Proper Treatment of Events*, Blackwell