

Anaphora and Direct Reference: Empirical Evidence from Pointing

Massimo Poesio
Università di Trento

Hannes Rieser
University of Bielefeld

Abstract

Empirical evidence from body measurements suggests that the referent of a demonstration is not directly specified, but obtained by applying a default inference rule to the region specified by the pointing cone. Building on this evidence we propose a unified theory of anaphoric and demonstrative uses in which accessibility is obtained via resource situations.

1 Introduction

The traditional semantics of demonstrative expressions is based on a sharp distinction between anaphoric reference and direct reference derived from Kaplan. Kaplan proposed that

‘[] each demonstrative, *d*, will be accompanied by a *demonstration*, δ , thus: $d[\delta]$. The character of a complete demonstrative is given by the semantic rule: In any context *c*, $d[\delta]$ is a directly referential term that designates *the demonstratum*, if any, of δ in *c*, and that otherwise designates nothing. Obvious adjustments are to be made to take into account any common noun phrase which accompanies or is built into the demonstrative.’ (Kaplan 1978, pp. 771-772).

Thus, for instance, demonstrative *This chair* in *This chair was hand-made by an artisan* accompanied by a pointing gesture to the chair (the demonstration) is interpreted as direct reference to the chair. By contrast, *This chair* in the text *Hannes bought a chair in the centre of Rovereto. This chair was hand-made by an artisan* is anaphoric. The two expressions have radically different interpretations.

This distinction has been challenged by semanticists such as Roberts (2002), as well as in corpus linguistics (Gundel et al, 1993); we will argue in this paper that it is also seriously challenged by empirical evidence about pointing. Modern body tracking methods make it possible to measure

with precision what a subject is pointing at. In a study combining experiments, statistical investigation, computer simulation and theoretical modelling techniques, Lücking, Pfeiffer and Rieser (2009) investigated the semantics and pragmatics of co-verbal pointing in dialogue. Lücking, Pfeiffer and Rieser established a semantic and two pragmatic hypotheses concerning the role of pointing in multi-modal expressions, and tested these with an annotated and rated corpus of Object Identification Games. The corpus was set up in experiments in which body movement tracking techniques were used to generate a space of pointing measurements. Statistical investigation and simulations showed that especially pointing to distal areas is not consistent with the semantic hypothesis. On the other end, the results can be predicted with high accuracy by hypothesizing a simple default inference extracting from the pointing gesture information sufficient to identify a referent uniquely. These results cast serious doubt on classical theories of the semantics-pragmatics interfaces insofar as they indicate that compositionality often presupposes pragmatically computed values.

In the paper we summarize the results of the Lücking *et al* study and formulate a unified hypothesis about the interpretation of demonstratives in terms of PTT (Poesio & Traum, 1997; Poesio & Rieser, submitted), a theory of the semantics and pragmatics of dialogue in which all actions in the discourse situation are explicitly represented and in which default inferences leading to their connection can be formulated.

2 A Brief Introduction to PTT

PTT (Poesio and Traum, 1997; Poesio & Muskens, 1997; Poesio & Rieser, submitted) is a theory of dialogue semantics and dialogue interpretation developed to explain how utterances are incrementally interpreted in dialogue, considering both their semantic impact and their impact on aspects of dialogue interaction traditionally considered as outside the scope of semantic theory, building on the work of Clark (1996) and on ideas from Situation Semantics (Barwise and

Perry, 1983; Cooper, 1996, Ginzburg, to appear). In this section we briefly discuss the two aspects of the theory that are relevant for the formulation of our unified hypothesis about demonstratives; for more details on PTT, including a complete fragment for German, see (Poesio & Rieser, submitted).

2.1 The common ground as a record of the discourse situation

PTT is an INFORMATION STATE theory of dialogue (Larsson & Traum, 2000; Stone, 2004; Ginzburg, to appear) in which the participants in a conversation maintain an information state about the conversation consisting of private information together with a conversational score including ‘grounded’ (Clark, 1996) and semi-public information. One respect in which PTT derives from Situation Semantics is hypothesis that the conversational score consists of a record of all actions performed during the conversation, i.e., what in Situation Semantics is called the DISCOURSE SITUATION (Barwise and Perry, 1983; Ginzburg, to appear). An ordinary conversation does not consist only of actions performed to assert or query a proposition, but also of actions whose function is to acquire, keep, or release a turn, to signal how the current utterance relates to what has been said before, or to acknowledge what has just been uttered. The discourse situation also contains information about non-verbal actions such as pointing.

Poesio and Traum (1997) argued that this view of the conversational score could be formalized using the tools already introduced in DRT (Kamp and Reyle, 1993)—specifically, in Muskens’s Compositional DRT (1996), because speech acts—CONVERSATIONAL EVENTS, in PTT terms—and non verbal actions are in many respects just like any other events, and because conversational events and their propositional contents can serve as the antecedents of anaphoric expressions. For instance, Poesio & Rieser (submitted) hypothesize that the two directives in (1) (an edited version of two turns from the Bielefeld ToyPlane Corpus) result in the update to the common ground in (2).¹

- (1) Inst: So jetzt nimmst Du eine orangene Schraube mit einem Schlitz

¹ We use the syntax from Muskens (1996) and his *is* operator. We use terms with the prefix *ce* (*ce1*, *ce2*, etc) for discourse referents denoting conversational events; terms with the prefix *u* for utterances; terms denoting other events will be indicated by the prefix *e*. We will indicate terms denoting states by the prefix *s*; all other terms will have prefixes *x*, *w*, *y*, and *z*.

So now you take a orange screw with a slit
 Cnst: Ja
 OK
 Inst: Und steckst Sie dadurch, von oben, daß also die drei festgeschraubt werden dann *and you put it through from above so that the three get fixed*

- (2) [K1.1, up1.1, ce1.1, K2.1, up2.1, ce2.1 |
 K1.1 *is* [e,x,x3] *screw*(x), *orange*(x),
slit(x3), *has*(x,x3),
e:grasp(Cnst, x)],
 up1.1: *utter*(Inst, “So jetzt nimmst Du ... “),
sem(up1.1) *is* K1.1,
 ce1.1: *directive*(Inst&Cnst,Cnst,K1.1),
generate(up1.1, ce1.1),
 K2.1 *is* [x6,e’,s,w,y] *x6 is x*,
e’:put-through(Cnst,x6,hole1),
w is wing1, *y is* fuselage1,
s: fastened(w,y)],
 up2.1: *utter*(Inst, “und steckst Sie ... “),
sem(up2.1) *is* K2.1,
 ce2.1: *directive*(Inst,Cnst, K2.1) ,
generate(up2.1, ce2.1)]

(2) records the occurrence of two conversational events, *ce1.1* and *ce1.2*, both of type **directive** (Matheson et al, 2000) whose propositional contents are separate DRSs specifying the interpretation of the two utterances in (1). The contents of conversational events are associated with propositional discourse referents K1.1 and K2.1 (discourse referents whose values are DRSs) as proposed in (Poesio and Muskens, 1997) and done, e.g., in SDRT (Asher and Lascarides, 2003). It is further assumed in PTT that dialogue acts are **generated** (Pollack, 1986) by locutionary acts (Austin, 1962), which we represent here as events of type **utter**.

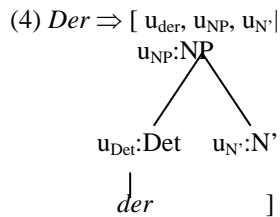
Non-verbal actions are also viewed in PTT as conversational events albeit of a different type. So for instance an act of pointing by agent DG would lead to the following update of both agents’ information state:

- (3) [pe1.1] pe1.1: *point*(DG, α)

where α is what DG is pointing at—determining experimentally what is α was the main question addressed by (Lücking, Pfeiffer and Rieser, 2009), as we will see.

It is assumed in PTT (Poesio, 1995) that the conversational score is incrementally updated whenever a verbal or non-verbal event is perceived. In particular, each word incrementally updates the discourse situation with a locutionary act of type **utter** and with syntactic expectations about the occurrence of more complex utterances as hypothesized in LTAG (Schabes et al, 1988). Thus, for instance, an utterance of definite article *der* results in the conversational score being updated with the occurrence of an utterance u_{der} of type Det (a *micro conversational event* (MCE) (Poe-

sio 1995a)) and with the expectation that this utterance will be part of an utterance a of a NP which will also include an utterance $u_{N'}$ of an N' . We will depict this update as follows:

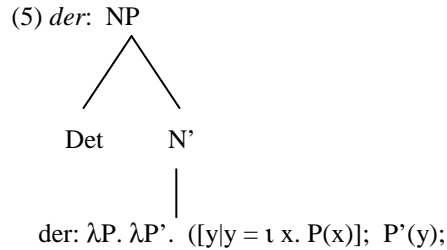


We further assume that MCEs have a (conventional) semantics associated to them, and that this semantics is the value of a **sem** function (in fact, a family of functions **sem**_[ι], **sem**_[π], etc.). We assume that the lexical semantics of words that update the discourse model and of anaphoric expressions is as proposed in Compositional DRT (Muskens, 1996), as discussed below, and that the semantics of phrasal utterances is obtained compositionally via defeasible inference rules that by default assign, for instance, to an utterance of an NP like u_{NP} above the conventional semantics **sem**(u_{NP}) resulting from the application of **sem**(u_{der}) to **sem**($u_{N'}$), but that can be overridden e.g., in the case of metonymy or as in the case of anaphoric expressions, as we will see below (Poesio & Traum 1997, Poesio to appear, Poesio & Rieser submitted).

2.2 Anaphora in PTT

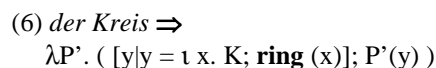
The current treatment of definites and anaphoric expressions in PTT (Poesio, to appear) is based on the ‘functional’ interpretation of definite NPs due to Loebner (1987) but has many points in common with the treatment proposed e.g., in (Chierchia, 1995). According to Loebner, what all definites have in common is that they are terms – i.e., *functions* that may take a different number of arguments, but all have a value of type e . Thus, for example, proper name *Jack* would have as translation the (0-argument) function $\iota x. (x = j)$, whereas the definite description *the pope* would have as translation the 1-argument function $\lambda s. \iota x. (x = \mathbf{pope}(s)(x))$, taking a situational or temporal argument s .

The Loebnerian treatment of definite descriptions is translated in the PTT framework by assigning to the definite article (e.g., German *der*) an elementary tree with the CDRT semantics below.



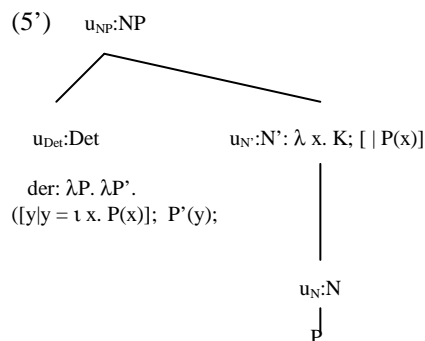
According to Loebner, a definite is licenced either because P is semantically functional, as in classical examples like *the king of France*, or because P is turned into a function by a modifier, as in *the first point to make is that...*, or because P is pragmatically *coerced* into a function by resolving it.

Standard DRT accessibility would predict that anaphoric interpretation in discourse situations is not possible: e.g., it would predict that the antecedent of *Sie* in (1), the screw, is not accessible. But Poesio (1993) proposed that what makes antecedents accessible in discourse situations is that definites uniformly receive their interpretation through a **resource situation** (Barwise and Perry, 1982; Cooper, 1996, Ginzburg, to appear). The resource situation hypothesis was recast in DRT terms in (Poesio 1994, Poesio & Muskens, 1997) by proposing that resource situations are contexts—DRSs—and that all anaphoric expressions contain an implicit variable over contexts, and it is this variable that supplies the value for the discourse referent. So for instance the NP *der Kreis* interpreted anaphorically would receive the following presuppositional interpretation:



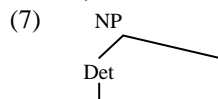
Where K is a resource situation where an object of type **ring** is particularly salient. (Note that K is used *presuppositionally*.)

The anaphoric interpretation in (6) is obtained through a *coercion process*—a defeasible semantic composition rule—that assigns to the N' in a definite construction as an interpretation a predicate $\lambda x. K; [| P(x)]$ that is pragmatically functional wrt a resource situation K , as in (5’):



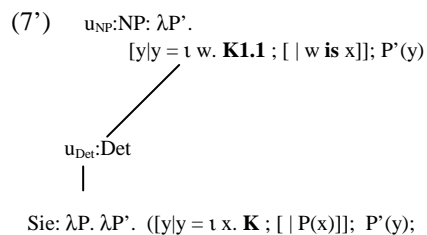
These coercion rules were called **Principles for Anchoring Resource Situations (PARS)** in (Poesio, 1993; Poesio, 1994). One such principle ruled anaphora, licensing the coercion above when the content K of a speech act is globally salient and contains an object of the right type. (Full specification of the principle omitted for reasons of space.) A second principle made parts of the visual scene salient as results of instructions that directed the attention to those parts of the scene. We will argue here that the evidence from Lücking *et al* suggests that pointing is another way for anchoring resource situations, thus providing a unified account of all types of definite reference, as already proposed by e.g., Roberts (2002) whose account, however, differs from ours in crucial respects.

It has often been argued that, syntactically, pronouns in English are like determiners. The translation proposed for pronouns such as *Sie* in (7) makes pronouns behave semantically like determiners, as well.



Sie: $\lambda P. \lambda P'. ([y|y = \iota x. K(x)]; P(y); P'(y))$

This translation is based on the idea that whereas the definite article may be licenced by a semantically functional, but non anaphoric, predicate, pronouns must always be licenced pragmatically—i.e., there must be some highly salient resource situation K containing a highly salient object. Furthermore, pronouns require a contextual property restricting the interpretation of the referent y : resolving a pronoun amounts to identifying such restriction. One obvious candidate is an identity property—i.e. a property of the form $\lambda w ([| w \text{ is } z])$ for z a discourse entity. According to the treatment just sketched, resolving *Sie* in (1) involves identifying $K1.1$ in (2) as resource situation and x as antecedent (i.e., applying the result to the identity property $\lambda w ([| w \text{ is } z])$), obtaining the following interpretation.



3 Experimental Evidence on Pointing

3.1 Semantic and Pragmatic Hypotheses on Pointing

Putting together assumptions by the early Wittgenstein, Davidson and Kaplan (1978), we can formulate the **Semantic Hypothesis** about pointing as follows:

(Sem) A demonstration [pointing] going together with a simple or a complex demonstrative in context c designates exactly one object, the object referred to in c .

The experimental literature in experimental pragmatics (Bangerter, 2004; Bangerter & Oppenheimer, 2006; Clark, 2003; Clark & Bangerter, 2004), however, leads to two rivalling hypotheses. The first one shifts the emphasis to inference to an object (*Strong Prag*); the second one deals with the focus of attention (*Weak Prag*) doing away with the notion of an object referred to altogether.

(Strong Prag) A demonstration triggers a perceptually based inference wrt a context c from the pointing device to the object referred to in c .

(Weak Prag) Demonstration shifts its addressee's attention towards a specific domain in a context c .

If one can show that (*Sem*) characterizes pointing behaviour in general, one does not need the pragmatics hypotheses, since pointing acts behave like constants. If one finds out that pointing success depends on contextual parameters, one has to resort to pragmatic hypotheses. Furthermore, if one finds evidence for (*Strong Prag*), one obviously has proved (*Weak Prag*), granted that one ties (*Strong Prag*) to intention and attention. Anyway, (*Weak Prag*) alone is not of much help, since it is too weak to distinguish pointing from focusing or emphasizing. For this reason we only concentrate on *StrongPrag* here.

3.2 Experimental Methods

In order to test the semantic and pragmatic hypotheses Lücking, Pfeiffer and Rieser (2009) conducted an empirical study using a so-called **Object identification game setting**. In this setting there are two participants, called *Description Giver* (DG) and *Object Identifier* (OI). DG and OI are set within the operational area of a marker-based optical tracking system with nine cameras (6DOF tracker, ART GmbH). The information delivered by the cameras provides positions and orientations of optical markers in an absolute coordinate system. Only the DG is

tracked by markers on arms, index fingers, hands, and head. Both OI and DG are located around a table (77.5×155.5 cm) with 32 parts of a *Lorentz Baufix* toy air-plane, the experimental domain.

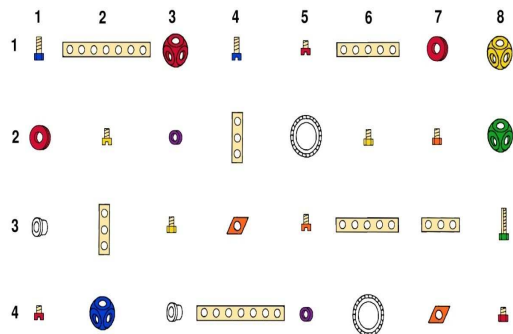


Figure 1: The experimental domain.

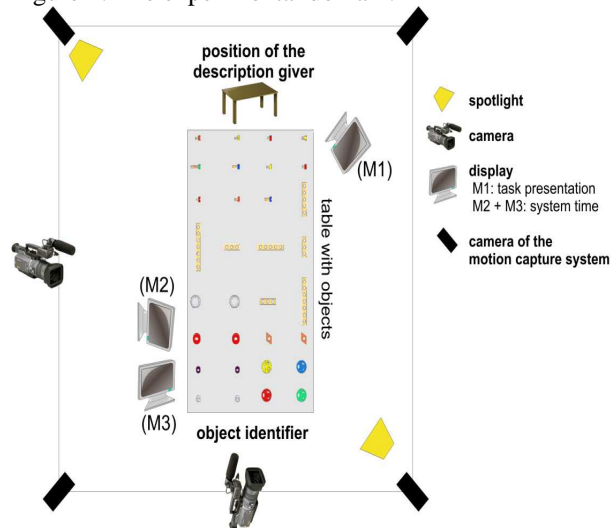


Figure 2: Setup of the setting within the interaction space of the motion capturing system. The interaction is observed by two video cameras and nine cameras of a motion capturing system.

The interaction between DG and OI was restricted to avoid negotiation processes. It consists of three steps: 1. Demonstration by DG (bimodal or only gestural); 2. Interpretation and identification by OI with a pointer (the referent remains in its place); 3. Verbal feedback by DG. The dialogues in the object identification games were of the following sort (original data):

- (8) DG: Der weiße Kreis da bei mir direkt auf der Linie, der weiße Kreis, der Reifen da.
The white circle near to me directly on the line, the white circle, the ring here.
 OI: [pointing].
 DG: Ja.
 OK.

3.3 Operationalization of the Hypotheses, Results and Analysis

The precise measurements of the motion capturing system provide us with the means to closely investigate pointing, reconstructing position and orientation of the index finger during each stroke. We also know the positions of the objects on the table. Thus Lücking, Pfeiffer and Rieser were able to project for each demonstration the beam from the index finger at the time of the stroke and compute whether the ray hits an object. It can be determined by the orientation of the index finger (index-finger-pointing, IFP) or, alternatively, by the direction of gaze, aiming at the target over the tip of the finger (gaze-finger-pointing, GFP).

Testing the (*Sem*) hypothesis on the pointing gesture means translating it in terms of predictions that can be measured using these methods. Lücking *et al* proposed the following:

Strict Operationalisation of the (*Sem*) hypothesis: *A pointing gesture refers to the object which is hit by a pointing-ray extending from the index-finger.*

If we calculate for each variant a pointing-ray originating in the index finger, oriented along the specific direction and intersect it with the table surface, we get a distribution of points around the object showing precision and accuracy of the pointing gesture (see Fig. 3).

Looking at Fig. 3, we see that pointing is fuzzy. In most of the demonstrations the projected ray fails the target. Reconsidering the semantic hypothesis in the context of the results shown in the bagplots of Fig. 3, a more relaxed conceptualization comes to mind which could compensate for the low precision of pointing but still allows us to sustain the (*Sem*) hypothesis. This leads to a relaxed operationalization of the (*Sem*) hypothesis using a pointing-cone to model the low precision of pointing:

Relaxed Operationalisation of the (*Sem*) hypothesis: *A pointing gesture refers to the object which is hit by a pointing-cone extending from the index-finger.*

However, the success rates were too low to provide a foundation for the weaker (*Sem*) hypothesis, leading finally Lücking *et al* to conclude that pointing is not a semantic referring device.

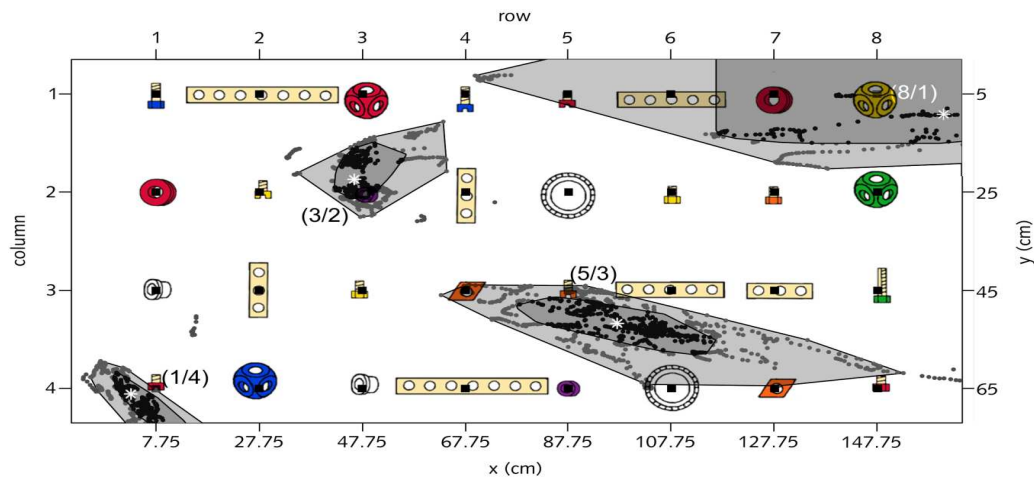


Fig. 3: The distribution of the objects on the table is overlaid with bagplots visualizing precision and accuracy of the pointing gestures for four selected objects (indicated by the pair of coordinates). The dots mark the intersections of a pointing ray with the table surface; a star indicates the mean position. Darker shading covers 50 percent and lighter shading 75 percent of the points. Obviously, most of the rays fail to hit the target object.

As stated, for semantics we would need a test providing a definite single object for every demonstration. This is different in pragmatics. Here we can use inference to choose among a set of possible referents selecting the most likely one intended by DG. Examining the (*Strong Prag*) hypothesis we only used motion capturing data. An example of inference process identifying one object among the objects in the pointing cone could be one that ranks the delimited objects according to their distance from the central axis of the pointing-cone. Lücking *et al* called this heuristics (INF):

- (INF) *An object is referred to by pointing only if*
- a) *the object is intersected by the pointing cone and*
 - b) *the distance of this object from the central axis of the cone is less than any other object's distance within this cone.*

Lücking *et al* further weakened their relaxed operationalisation for the (*Sem*) hypothesis and allowed several objects to lie within the pointing-cone as long as the intended target object can be singled out from the set of objects delimited *via* a subsequent inference. So they arrived at the following:

Operationalisation of the (*Strong Prag*) hypothesis: *A pointing gesture refers to the one object selected by an appropriate inference from the set of objects covered by a pointing-cone extending from the index-finger.*

(In other words, the object demonstrated is the one nearest to the axis of the pointing cone where a) and b) are considered to be necessary conditions.)

This weighting heuristics succeeds in 96 percent of the cases when using Index-Finger-Pointing and in 92 percent of the cases when using Gaze-Finger-Pointing. These results are mainly due to the weighting heuristics and not to a clear-cut cone intersection. We take these figures as strong evidence that (*Strong Prag*) holds, i.e., that the referent in demonstrative uses is arrived at via a pragmatic inference process which, however, is not infallible (i.e., it is a defeasible inference)

4 A Unified Account of Anaphoric And Demonstrative Uses

If it is true that the referents of demonstratives are obtained through an inference like (INF), then there is no need to stipulate that demonstrative phrases like *this chair* are ambiguous between an anaphoric and a direct reference use: the translation of definites proposed in 2.2 can serve as the lexical translation for definites like *der weisse Kreis* both when used anaphorically and when used demonstratively in (8).

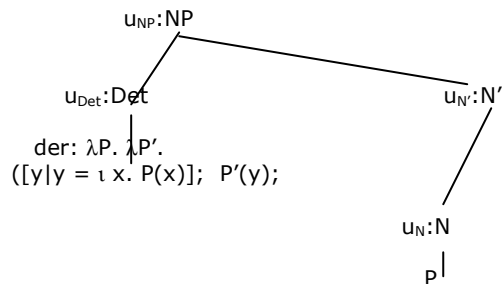
Assuming that the visual scene is a resource situation K_{visual} as proposed in (Poesio, 1993), then the results by Lücking, Pfeiffer and Rieser (2009) suggest that an act of pointing identifies a subset of this situation K_{pointing} the set of objects in the pointing cone.

(3') [pe1.1] pe1.1:point(DG, K_{pointing})

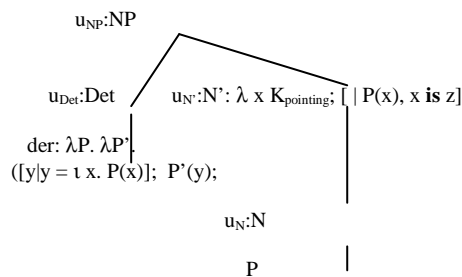
INF is a defeasible inference rule analogous to the Principles for Anchoring Resource Situations proposed in (Poesio, 1993), except that it coerces the interpretation of the nominal predicate to be identical with the object z in K_{pointing} which is the closest object to the central axis of the cone:

Principle for Anchoring Resource Situations via Pointing

If u_{NP} is a micro-conversational event with structure



u_{NP} is cotermporal with pointing gesture $pe1.1:\text{point}(DG, K_{\text{pointing}})$, and $z \in K_{\text{pointing}}$ is the closest object to the pointing axis of the cone, then defeasibly coerce the interpretation of N' to $\lambda x K_{\text{pointing}}; [| P(x), x \text{ is } z]$:



5 Related Work

Roberts (2002) aims at a unified account of several types of demonstratives, pronominal and descriptive, accompanied by canonical demonstrations or textual deixis and discourse deixis. The following quotation sums up her approach: *The heart of this proposal is the claim that a demonstrative NP conventionally presupposes that a familiar discourse referent for the demonstratum of its associated demonstration is the same as the discourse referent which satisfies the NP's familiarity presupposition.* (p. 315) There are similarities between the PTT-account of demonstrative definites and the one presented in Roberts: for example, the hypothesis that the definiteness information is presuppositional (p. 312). The difference between her account and the one presented here is that here notions like demon-

stration, demonstratum, pointing, direction pointed at, context, salience, proximity and non-proximity are part of the *explicandum* for which the experimental situation, including body tracking devices serves as an *explicans*. So little is taken for granted and notions are backed up by rigid measurements. Similarities and differences would merit a more thorough discussion.

6 Conclusions

Modern experimental techniques are beginning to make it possible to empirically test fine-grained semantic hypotheses. We argued in this paper that in particular evidence from body measurements can be used to evaluate the extent to which demonstrations uniquely identify the referent of a demonstrative. The evidence is that the identification requires a pragmatic inference process. The next step will be to revisit other linguistic evidence for direct reference at the light of these data.

Acknowledgments

We are grateful to Andy Lücking and Thies Pfeiffer for letting us use material from an unpublished paper. This research was in part supported by the University of Trento, Centre for Mind / Brain Sciences, and by the University of Bielefeld.

References

- Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. CUP
- Austin, J. L. (1962). *How to Do Things with Words*. Oxford, OUP
- Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, 15(6):415–419.
- Bangerter, A. and Oppenheimer, D. M. (2006). Accuracy in detecting referents of pointing gestures unaccompanied by language. *Gesture*, 6(1):85–102.
- Barwise, Jon & Perry, John (1983). *Situations and attitudes*. MIT Press.
- Butterworth, G. and Itakura, S. (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology*, 18:25–50.
- Chierchia, G. (1995). *Dynamics of Meaning*. Univ. of Chicago Press.
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Clark, H. H. (2003). Pointing and placing. In Kita, S., editor, *Pointing*, chapter 10, pages 243–269. Lawrence Erlbaum.
- Clark, H. H. and Bangerter, A. (2004). Changing ideas about reference. In Noveck, I. A. and Sperber, D., editors, *Experimental Pragmatics*, pages 25–49. Palgrave Macmillan, New York.
- Cooper, R. (1996). The role of situations in generalized quantifiers. In Lappin, S. (ed), *Handbook of Contemporary Semantic Theory*, Blackwell.
- Ginzburg, J.. To Appear. *Semantics for Conversation*. CSLI Publications
- Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Dordrecht: Kluwer.
- Kaplan, D. (1978). On the logic of demonstratives. *J. of Philosophical Logic*, 8:81–98.
- Larsson, S. and Traum, D. (2000). Information state and dialogue management in the TRINDI Dialogue Move Engine Toolkit. *Natural Language Engineering*, v.6, pp. 323-340.
- Loebner, S. (1987) Definites. *Journal of Semantics*, 4, 279-326
- Lücking, A., Pfeiffer, Th., and Rieser, H. (2009). Pointing and Reference Reconsidered (submitted)
- Matheson, C., M. Poesio, and D. Traum, (2000). Modeling Grounding and discourse obligations using update rules, *Proc. Of the NAACL*.
- Muskens, R. (1996). Combining Montague Semantics and Discourse Representation. *Linguistics and Philosophy* 19, 143-186
- Poesio, M. (1993). A Situation-Theoretic Formalization of Definite Description Interpretation. In P. Aczel, D. Israel, Y. Katagiri, and S. Peters, (editors), *Situation Theory and its Applications*, vol. 3, p.339--374.
- Poesio, M. (1994). *Discourse Interpretation and the Scope of Operators*. PhD dissertation, University of Rochester.
- Poesio, M. (1995). "A Model of Conversation Processing Based on Micro Conversational Events," In *Proceedings of the Cognitive Science Society*.
- Poesio, M. to appear. *Incrementality and underspecification in semantic interpretation*. CSLI Publications.
- Poesio, M. and Muskens, R. (1997). "The Dynamics of Discourse Situations", *Proc. 11th Amsterdam Colloquium*, 247-252.
- Poesio, M. and Rieser, H., submitted. *Completions, coordination and alignment in dialogue*.
- Poesio, M. and Traum, D. (1997). Conversational Actions and Discourse Situations, *Computational Intelligence*, v. 13, n.3, pp.1- 45
- Pollack, M. (1986), *Inferring Domain Plans in Question-Answering*. PhD, Univ. of Pennsylvania.
- Roberts, C. (2002). Demonstratives as definites. In K. van Deemter & R. Kibble (eds.) *Information Sharing*. CSLI Press, Stanford, 89-196.
- Schabes, Y., Abeillé, A., & Joshi, A.K. (1988). New parsing strategies for tree adjoining grammars. In *Proceedings of COLING*, pp. 578– 583. Budapest, August.
- Stone, M. (2004) Intention, interpretation and the computational structure of language. *Cognitive Science* 28(5):781-809.