Constraints on Linguistic Coreference: Structural vs. Pragmatic Factors

Frank Keller
keller@coli.uni-sb.de
Computational Linguistics, Saarland University
PO Box 15 11 50, Building 17.1
66041 Saarbrücken, Germany

Ash Asudeh
asudeh@csli.stanford.edu
Department of Linguistics, Stanford University
Margaret Jacks Hall, Building 460
Stanford CA 94305-2150, USA

Abstract

Binding theory is the component of grammar that regulates the interpretation of noun phrases. Certain syntactic configurations involving picture noun phrases (PNPs) are problematic for the standard formulation of binding theory, which has prompted competing proposals for revisions of the theory. Some authors have proposed an account based on structural constraints, while others have argued that anaphors in PNPs are exempt from binding theory, but subject to pragmatic restrictions. In this paper, we present an experimental study that aims to resolve this dispute. The results show that structural factors govern the binding possibilities in PNPs, while pragmatic factors play only a limited role. However, the structural factors identified differ from the ones standardly assumed.

Introduction

Linguistic Intuitions The data on which linguists base their theories typically consist of grammaticality judgments, i.e., intuitive judgments of the well-formedness of utterances in a given language. When a linguist obtains a grammaticality judgment, he or she performs a small experiment on a native speaker; the resulting data are behavioral data in the same way as other measurements of linguistic performance (e.g., the reaction time data used in psycholinguistics). However, in contrast to experimental psychologists, linguists are generally not concerned with methodological issues, and typically none of the standard experimental controls are imposed in collecting data for linguistic theory. As Schütze's (1996) recent work on empirical issues in linguistics demonstrates, such methodological negligence can seriously compromise the data obtained. Schütze (1996) argues for a more reliable mode of data elicitation in linguistics, based on standard methods from experimental psychology.

Recently, Bard, Robertson, and Sorace (1996) and Cowart (1997) demonstrated how the experimental paradigm of magnitude estimation (ME) makes it possible to address problems such as the ones raised by Schütze. ME is an experimental technique standardly used in psychophysics to measure judgments of sensory stimuli (Stevens, 1975). It requires subjects to estimate the magnitude of physical stimuli by assigning numerical values proportional to the stimulus magnitude they perceive. Highly stable judgments can be achieved for a whole range of sensory modalities, such as brightness, loudness, or tactile stimulation. Bard et al. (1996) demonstrated that linguistic judgments can be elicited in the same way as judgments of sensory stimuli, and that ME can yield reliable and fine-grained measurements of linguistic intuitions.

The present paper applies the ME methodology to a longstanding dispute in linguistic theory, viz., the binding theoretic status of picture noun phrases (PNPs). Binding in PNPs has generated considerable interest in the literature, and has prompted a number of revisions of standard binding theory. However, there is considerable disagreement on both the relevant data (i.e., coreference judgments for PNPs) and on the theoretical conclusions to be drawn from these data. In this paper, we demonstrate how the use of experimentally elicited coreference judgments can resolve such theoretical disputes.

Binding Theory and Exempt Anaphors Binding theory (BT) is the component of grammar that regulates the interpretation of noun phrases (NPs). Three types of noun phrases are generally distinguished: (a) full NPs (e.g., *Hanna*, *the woman*, *every woman*), (b) pronouns (e.g., *he*, *her*), and (c) anaphors (e.g., *herself*, *each other*). The task of BT is to determine which noun phrases in a given syntactic domain can be *coreferential*, i.e., refer to the same individual. Coreference is normally indicated with subscripts:

(1) a. Hanna$_i$ admires *her$_i$/herself$_i$.
b. Hanna$_i$ thinks that Peter admires her$_i$/*herself$_i$.

In example (1a), the proper name *Hanna* and the pronoun *her* cannot refer to the same person, i.e., they cannot be coreferential (as indicated by the '*'). The pronoun cannot be *bound* by the proper name. In (1b), on the other hand, *Hanna* is a potential binder for *her*, i.e., coreference is possible. The possibilities for anaphor binding are exactly reversed; *Hanna* must bind (i.e., corefer with) *herself* in (1a), but cannot do so in (1b).

There are distinct structural conditions that determine the binding possibilities for the different kinds of NPs. Principle C of BT deals with the binding requirement for full NPs, and will not concern us here. Principle A captures the binding requirements for anaphors; in early formulations, it states that an anaphor has to be bound within a certain local domain (Chomsky, 1981). The local domain is defined using c-command, a structural notion defined on trees. Principle B, on the other hand, states that pronouns cannot be bound within the local domain. It follows that anaphors and pronouns are predicted to be in complementary distribution, i.e., anaphors can be bound where pronouns cannot be bound, and vice versa.

It was subsequently observed that this complementarity breaks down in certain structures. A case that has generated much theoretical discussion is PNPs, where anaphors and pronouns are equally acceptable:

(2) Hanna$_i$ found a picture of her$_i$/herself$_i$.

There is also the further complication that in PNPs with possessors (3) and in PNPs that are arguments of certain verbs (4) the complementarity between pronouns and anaphors seems to resurface:

(3) Peter$_i$ found Hanna$_j$'s picture of *her$_j$/herself$_j$.
(4) Hanna$_i$ took a picture of *her$_i$/herself$_i$.

Note that (4) is meant with the sense of *take* as in creating a photograph, not as in physically removing a picture. A number of authors have argued for a revised version of BT based on data such as (2), (3), and (4). Chomsky (1986) restates BT such that there is an asymmetry between pronouns and anaphors in certain contexts, including PNPs without possessors. For (4), Chomsky and Lasnik (1995) propose that there is a covert possessor. With these revisions, the predicted pattern of data is exactly as in (2)–(4). We will refer to this approach as the *structural account* of binding in PNPs.

Proceedings of the 23rd Annual Conference of the Cognitive Science Society. Mahawah, NJ: Lawrence Erlbaum, 2001.

Some more recent work, however, has proposed a *pragmatic* account of the PNP data in (2)–(4) (e.g., Kuno, 1987; Pollard & Sag, 1994; Reinhart & Reuland, 1993). These authors have observed that in certain configurations anaphors are *exempt* from BT. One such configuration is PNPs without possessors, as in (2) and (4). According to this view, the anaphor in (2) is not subject to Principle A, but is rather governed by pragmatic constraints, where relevant factors include referentiality, definiteness, and aspect. It is important to note that even the versions of BT that postulate exempt anaphors still maintain that Principle A holds of anaphors in PNPs when there is an overt possessor: although the anaphors in (2) and (4) are exempt, the anaphor in (3) is subject to BT.

The present study attempts to clarify the empirical status of exempt anaphors. We present the results of an experiment that tests the influence of both structural and pragmatic factors on coreference in PNPs. This experiment uses the magnitude estimation (ME) paradigm to establish the coreference intuitions of linguistically naive subjects. (For other studies demonstrating the usefulness of experimental data in clarifying BT facts, see Cowart, 1997; Gordon & Hendrick, 1997.)

Before we discuss the results of this experiment, we present a control study designed to validate our experimental paradigm. To our knowledge, ME has never been applied to coreference judgments, hence we must show that its results are consistent with the theoretical literature and replicate previous experimental data.

Experiment 1: Control Study

The control study was designed as a replication of Experiment 3 of the study of coreference by Gordon and Hendrick (1997). It investigated basic effects of Principles A, B, and C of BT. Eight different binding configurations were tested, three of which occurred either with or without c-command (see Chomsky, 1981, for details on c-command). Table 1 lists the binding configuration tested by Gordon and Hendrick (1997). It also summarizes the predictions of standard BT for these configurations, and gives example stimuli.

Predictions

Our hypothesis is that ME generates valid coreference judgments. We therefore predict that the same significant effects as in Gordon and Hendrick's (1997) original study will be present, even though our replication used an ME task instead of the ordinal judgment task employed by Gordon and Hendrick (1997). Another difference is that we conducted our experiment over the World Wide Web, while Gordon and Hendrick (1997) administered a conventional questionnaire. The web-based methodology entails differences in sampling and experimental procedure, which increases the need for a validation study.

Method

Subjects Fifteen participants were recruited over the Internet by postings to newsgroups and mailing lists. All participants were self-reported native speakers of English and naive to syntactic theory.

Materials Following Gordon and Hendrick (1997), the design contained one factor, viz., binding configuration (*Ana*) with eight levels. Three lexicalizations were used; one was the original lexicalization used by Gordon and Hendrick (1997), the other two were new lexicalizations, designed in analogy with the original one. This resulted in a set of 24 items (see Table 1 for sample stimuli).

Procedure The method used was ME as proposed by Stevens (1975) for psychophysics and extended to linguistic stimuli by Bard et al. (1996) and Cowart (1997).

Subjects first saw a set of instructions that explained the concept of numerical ME using line length. Subjects were instructed to make length estimates relative to the first line they would see, the reference line. They were told to give the reference line an arbitrary number, and then assign a number to each following line so that it represented how long the line was in proportion to the reference line. Several example lines and corresponding numerical estimates were provided to illustrate the concept of proportionality. Then subjects were told that linguistic acceptability could be judged in the same way as line length, and that this experiment required them to judge the acceptability of coreference. Following Gordon and Hendrick (1997), this was defined as follows: 'Your task is to judge how acceptable each sentence is by assigning a number to it. By acceptability we mean the following: Every sentence will contain two expressions in ALL CAPITALS. A sentence is acceptable if these two expressions can refer to the same person.' The task was illustrated by examples.

After reading the instructions, subjects took part in a training phase designed to familiarize them with the task. In the training phase, subjects were ask to use ME to judge the length of a set of lines. Then, a set of practice items (similar to the experimental items) were administered to familiarize subjects with applying ME to linguistic stimuli. Finally, subjects had to judge the experimental items. Each subject judged all 24 experimental stimuli and a set of 24 fillers, i.e., a total of 48 items.

The experiment was conducted over the web using WebExp 2.1 (Keller, Corley, Corley, Konieczny, & Todirascu, 1998), an interactive software package for web-based psycholinguistic experimentation. Keller and Alexopoulou (2001) present a detailed discussion of the safeguards that WebExp puts in place to ensure the authenticity and validity of the data collected, and also present a validation study comparing web-based and lab-based judgment data (for a WebExp validation study using sentence completion data, see Corley & Scheepers, in press).

Results

The data were normalized by dividing each numeric judgment by the modulus value that the subject had assigned to the reference sentence. This operation creates a common scale for all subjects. Then the data were transformed by taking the decadic logarithm. This transformation ensures that the judgments are normally distributed and is standard practice for ME data (Bard et al., 1996). All analyses and figures are based on normalized, log-transformed judgments.

The average judgments for the different conditions are graphed in Figure 1 for the original study and in Figure 2 for our replication. Visual inspection shows that the replication experiment produces the same acceptability pattern for each of the binding configurations.

This was confirmed by the statistical analyses. Gordon and Hendrick (1997) report a significant main effect of binding configuration (*Ana*), which was also present in our data ($F_1(7,98) = 17.561$, $p < .0005$; $F_2(7,14) = 295.262$, $p < .0005$). They also found that the acceptability of the name-anaphor configuration increased under c-command, which was replicated in our data ($F_1(1,14) = 17.057$, $p = .001$; $F_2(1,2) = 2389.474$, $p < .0005$). Another finding was that c-command significantly reduces the acceptability of coreference name-pronoun configurations. This effect was also present in the replica-

Table 1: Sample stimuli and predictions from Gordon and Hendrick (1997), Experiment 3

NP_1	NP_2	c-command	sample sentence	prediction
name	pronoun	no	(i) **Joan's** father respects **her.**	grammatical
pronoun	name	no	(ii) **Her** father respects **Joan.**	grammatical
name	name	no	(iii) **Joan's** father respects **Joan.**	grammatical
pronoun	anaphor	no	(iv) **Her** father respects **herself.**	ungrammatical
name	anaphor	no	(v) **Joan's** father respects **herself.**	ungrammatical
name	pronoun	yes	(vi) **Joan** respects **her.**	ungrammatical
pronoun	name	yes	(vii) **She** respects **Joan.**	ungrammatical
name	anaphor	yes	(viii) **Joan** respects **herself.**	grammatical

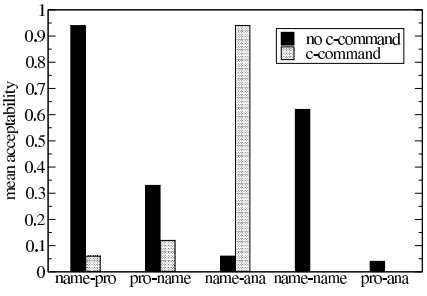

Figure 1: Original data from Gordon and Hendrick (1997), Experiment 3

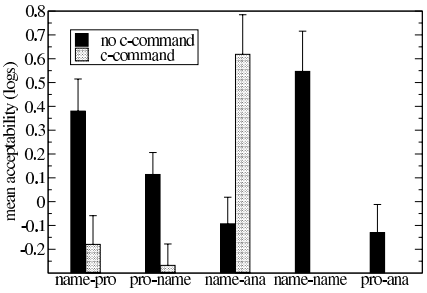

Figure 2: Replication of Gordon and Hendrick (1997), Experiment 3

tion ($F_1(1,14) = 21.818$, $p < .0005$; $F_2(1,2) = 315.306$, $p = .003$). An effect of c-command on the acceptability of pronoun-name configurations was also found both in the original data set and in our replication ($F_1(1,14) = 25.949$, $p < .0005$; $F_2(1,2) = 181.980$, $p = .005$).[1] Finally, a comparison of the name-pronoun and the name-name configurations showed that names are favored as antecedents ($F_1(1,14) = 13.770$, $p < .002$; $F_2(1,2) = 192.301$, $p = .005$), in line with what Gordon and Hendrick (1997) found.

To further compare the results of the original experiment and our validation study, we conducted a correlation analysis comparing the mean judgments for each cell in the experiment. A high correlation coefficient was obtained ($r_1 = .9198$, $p = .001$, $N = 8$). (No by-item correlation coefficient could be computed as Gordon and Hendrick (1997) fail to report by-item analyses.)

[1]Note that standard BT fails to predict the reduced acceptability of configuration (ii). A possible explanation might be that this configuration involves cataphoric reference (i.e., the pronoun refers forwards instead of backwards).

Discussion

In this study, we used ME to replicate a published experiment on coreference judgments that used a conventional ordinal scale (Gordon and Hendrick's (1997) Experiment 3). We obtained the same significant effects as in the original and a high correlation with the original data set, which amounts to a full replication of the original study.

This result indicates that the ME paradigm is suitable for investigating judgments of linguistic coreference, which are vital for testing claims from BT. Previous uses of ME were limited to grammaticality judgments (Bard et al., 1996; Cowart, 1997). The successful replication also reassures us that psycholinguistic data collected over the web yield results comparable to data generated by a conventional lab-based methodology, in line with previous findings by Keller and Alexopoulou (2001) and Corley and Scheepers (in press).

Finally, the present experiment allows us to establish a baseline for further experiments on linguistic coreference. It encompassed only clear-cut cases of coreference that are uncontroversial in the binding theoretical literature. It is important to establish the validity of our methodology for such clear-cut cases before moving on to investigate more controversial issues such as binding in PNPs, where the theoretical and empirical claims in the syntactic literature differ widely. Binding in PNPs is the subject of the next experiment.

Experiment 2: Structural and Pragmatic Factors in Coreference

Based on the results from the control experiment, we carried out an experimental study investigating the factors that determine coreference in PNPs in English. The aim of this experiment was to provide reliable experimental data that settles the longstanding dispute about the binding theoretical status of PNPs. In particular, we tested the claim that PNPs are exempt from BT, and hence their coreference options are governed by pragmatic, rather than structural factors.

Structural Factors The current experiment tested the influence of structural factors on binding in PNPs by comparing the behavior of anaphors and pronouns in six configurations, listed in Table 2. Two structural factors were tested.

Firstly, the position of the binder, which can either be the subject of the matrix clause (as in configurations (i)–(iv) in Table 2), or the possessor of the PNP (as in configurations (v) and (vi) in Table 2). Secondly, the absence of a possessor (as in configurations (i) and (ii)), or its presence (as in configurations (iii)–(vi) in Table 2). The experiment contained three subdesigns, which tested the configurations (i) and (ii), (iii) and (iv), and (v) and (vi), respectively.

Table 2: Sample stimuli and predictions for Experiment 2

NP_1	NP_2	subject	possessor	sample sentence	prediction
name	pronoun	yes	no	(i) **Hanna** found a picture of **her**.	grammatical
name	anaphor	yes	no	(ii) **Hanna** found a picture of **herself**.	grammatical
name	pronoun	yes	yes	(iii) **Hanna** found Peter's picture of **her**.	grammatical
name	anaphor	yes	yes	(iv) **Hanna** found Peter's picture of **herself**.	ungrammatical
name	pronoun	no	yes	(v) Hanna found **Peter's** picture of **him**.	ungrammatical
name	anaphor	no	yes	(vi) Hanna found **Peter's** picture of **himself**.	grammatical

Pragmatic Factors The second aim of the present experiment was to investigate the influence of pragmatic factors on coreference in PNPs. Such factors have received much attention in the theoretical literature. However, no quantitative studies have been conducted to determine to what extent these factors influence coreference, and how they interact with structural factors.

Three pragmatic factors were investigated. The first one is definiteness of the PNP. As an example of definiteness consider the minimal pair in (5): the PNP in (5a) is indefinite and the one in (5b) is definite.

(5) a. Hanna$_i$ found a picture of her$_i$/herself$_i$.
b. Hanna$_i$ found the picture of her$_i$/herself$_i$.

The second factor is the aspectual class of the matrix verb, illustrated in example (6): *find* and *lose* are examples of achievement verbs, while *take* and *destroy* are accomplishment verbs; *find* and *take* are [+existence], i.e., they presuppose the existence of their object, while *lose* and *destroy* are [−existence], i.e., they do not carry this presupposition.

(6) a. Hanna$_i$ found a picture of her$_i$/herself$_i$.
b. Hanna$_i$ lost a picture of her$_i$/herself$_i$.
c. Hanna$_i$ took a picture of her$_i$/herself$_i$.
d. Hanna$_i$ destroyed a picture of her$_i$/herself$_i$.

Third, we tested the influence of the referentiality of the binder, as illustrated in (7):

(7) a. Hanna$_i$ found Peter's picture of her$_i$/herself$_i$.
b. The woman$_i$ found Peter's picture of her$_i$/herself$_i$.
c. Each woman$_i$ found Peter's picture of her$_i$/herself$_i$.

The pragmatic factors were included in the three subdesigns of the present experiment. Definiteness and verb class were included in the first subdesign, and referentiality was included in the second and third subdesigns.

Predictions

Based on the theoretical literature (see Introduction), we predict that anaphors in PNPs are exempt from local binding (i.e., binding within the PNP), *unless* the PNP has a possessor, in which case the anaphor must be bound by the possessor (see examples (2) and (3)). We also predict that pronouns must be locally free from a possessor, if there is one. Table 2 lists the configurations and the associated predictions, together with example stimuli. Note that we expect that the relative acceptability of pronouns and anaphors is the same in configurations (i), (ii) and (iii). Configurations (iv) and (v) are predicted to be unacceptable, while (vi) is predicted to be acceptable. These constructions differ in terms of their syntactic structure (antecedent is the subject or the possessor; possessor is present or not). We expect to find no main effect of binding configuration for (i) versus (ii), but for pairs (iii)/(iv) and (v)/(vi) we expect binding configuration to have a significant main effect.

If the pragmatic approach to binding in PNPs is correct, then we also expect that the pragmatic factors verb class, definiteness, and referentiality have an effect on coreference. The underlying theoretical assumption is that coreference for exempt NPs is governed by pragmatics, rather than by structural principles. Hence we predict an interaction of binding configuration with verb class and definiteness in the first subexperiment, and an interaction of binding configuration with referentiality in the second and third subexperiments.

Method

Subjects Fifty-two native speakers of English volunteered to participate. All participants were naive to syntactic theory.

Materials The experimental materials included three subdesigns. The first subdesign investigated binding configurations (i) and (ii): name-pronoun and name-anaphor with the antecedent in the subject and without a possessor. The second subdesign compared binding configurations (iii) and (iv): name-pronoun and name-anaphor with the antecedent in the subject and a possessor in the PNP. The third subdesign dealt with configurations (v) and (vi): name-pronoun and name-anaphor with the antecedent as the possessor of the PNP.

This means that in each of the three subdesigns the factor binding configuration (*Ana*) had two levels: name-pronoun or name-anaphor. In the first subdesign, this factor was crossed with *Def* and *Verb*, representing the two pragmatic factors definiteness of the PNP and aspectual class of the main verb. *Def* had two levels (definite, indefinite, see (5)), *Verb* had three levels (achievement [+existence], accomplishment [+existence], accomplishment [−existence] (see (6a), (6c), (6d)). This yielded a total of $Ana \times Def \times Verb = 2 \times 2 \times 3 = 12$ cells for the first subdesign.

In the second and third subdesigns, the structural factor *Ana* was crossed with the pragmatic factor referentiality (*Ref*), which had three levels (proper name, definite NP, quantified NP, see (7)). The second and third subdesigns hence had $Ana \times Ref = 2 \times 3 = 6$ cells each.

The three subdesigns taken together had a total of 24 cells. Four lexicalizations were used for each cell, resulting in a total of 96 stimuli. A set of 24 fillers was used, covering the whole acceptability range.

Procedure The experimental procedure was the same as in Experiment 1. The stimulus set was divided into four subsets of 24 stimuli by placing the items in a Latin square. Each subject judged one of these subsets and 24 fillers, i.e., a total of 48 items.

Results

The data were preprocessed as in Experiment 1. Separate ANOVAs were conducted for each subexperiment.

Structural Factors In the first subexperiment (binding configurations (i) and (ii)), we found a large and highly significant main effect of *Ana* ($F_1(1,51) = 137.471$, $p < .0005$; $F_2(1,3) = 105.005$, $p = .002$). Anaphors (mean = .6702) were more acceptable than pronouns (mean =

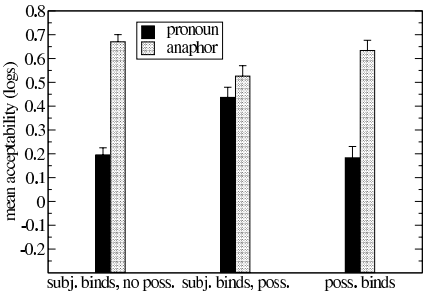

Figure 3: Structural effects on coreference judgments for binding in PNPs

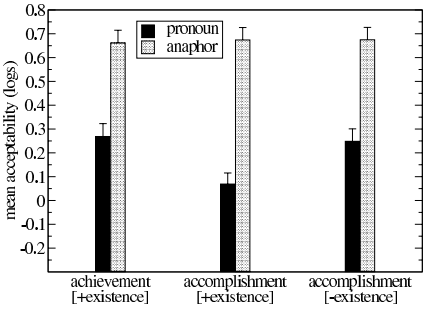

Figure 4: Effect of verb class on coreference judgments (subject binds, no possessor)

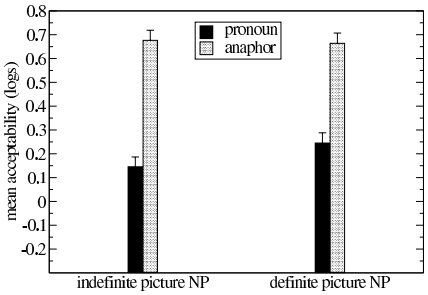

Figure 5: Effect of definiteness on coreference judgments (subject binds, no possessor)

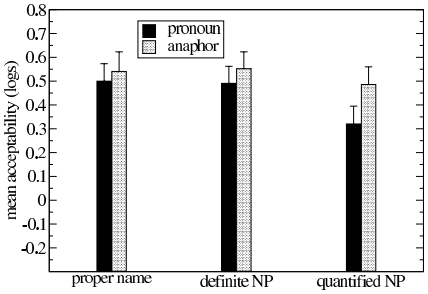

Figure 6: Effect of referentiality on coreference judgments (subject binds, possessor)

.1954). In the second subexperiment (binding configurations (iii) and (iv)), *Ana* failed to reach significance: both anaphors (mean = .5262) and pronouns (mean = .4369) were equally acceptable. In the third subexperiment (binding configurations (v) and (vi)), again a main effect of *Ana* was present ($F_1(1,51) = 101.632$, $p < .0005$; $F_2(1,3) = 34.677$, $p = .010$). Anaphors (mean = .6338) were more acceptable than pronouns (mean = .1832). The coreference judgments for the six binding configurations (see Table 2) are graphed in Figure 3.

Pragmatic Factors The ANOVA for the first subexperiment also revealed a significant interaction of *Verb* and *Ana* ($F_1(2,102) = 11.275$, $p < .0005$; $F_2(2,6) = 6.193$, $p = .035$). This interaction is graphed in Figure 4, showing a decrease in the acceptability of pronouns for [+existence] accomplishment verbs. An interaction of *Def* and *Ana* was also found, which however was significant by subjects only ($F_1(1,51) = 11.849$, $p = .001$; $F_2(1,3) = 2.168$, $p = .237$). Figure 5 shows that the acceptability of pronouns is increased for definite PNPs.

The ANOVA for the second subexperiment showed an interaction of *Ref* and *Ana*, significant by subjects only ($F_1(2,102) = 3.979$, $p = .049$; $F_2(2,6) = 2.745$, $p = .142$). This interaction is depicted in Figure 6, showing a decrease in the acceptability of pronouns if the antecedent is a quantified NP. No *Ref*/*Ana* interaction was present in the third subexperiment (see Figure 7).

Discussion

The theoretical predictions for the acceptability of the stimuli are listed in Table 2. Theory also predicts that anaphors are exempt from BT in configuration (i), and that structural factors should fail to have an influence on the acceptability of coreference for these structures. Contrary to this prediction, the present experiment revealed a significant influence of structural factors, although not in a way that any existing account predicts. Four major results were obtained.

In cases where the antecedent is in the subject and there is no possessor in the PNP (configurations (i) and (ii), see Table 2), structural and pragmatic binding theories alike predict that pronouns are fully acceptable and that pronouns and anaphors are equally acceptable. Our first and second major results are the falsification of both these predictions. Pronouns were significantly less acceptable than anaphors (see Figure 3). A comparison with standard cases of BT tested in Experiment 1 (see Figure 2) indicates that anaphors are fully acceptable in this configuration, while pronouns are of intermediate acceptability (but not fully unacceptable compared to, e.g., name-pronoun configurations with c-command).

Configurations (iii) and (iv), where the antecedent is in the subject, but there is a possessor, demonstrate our third major result. Here BT falsely predicts that anaphors are fully unacceptable. Note also that these anaphors

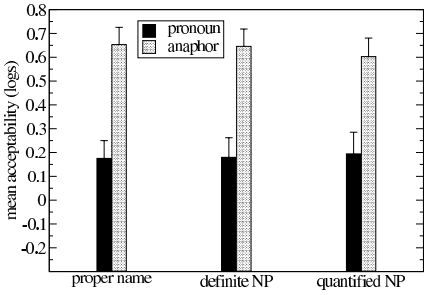

Figure 7: Effect of referentiality on coreference judgments (possessor binds)

are not exempt according to the pragmatic versions of BT, as there is a possessor. We found that pronouns and anaphors are both highly acceptable; no significant acceptability difference could be detected (see Figure 3). In other words, contrary to all that has been written in the syntactic literature, anaphors can be bound by the subject even in PNPs with possessors.

Our fourth result concerns the third structure we investigated (configurations (v) and (vi)), where the antecedent is the possessor in the PNP rather than the subject. We found the same behavior as in configurations (i) and (ii): the anaphors were fully acceptable in this configuration, while pronouns were significantly less acceptable, but not completely unacceptable compared to the configurations investigated in Experiment 1 (see Figure 2). The prediction for a PNP with a possessor is that a pronoun bound by the possessor is completely ungrammatical and that a pronoun bound by the subject is completely grammatical. This prediction was not supported by our results. We found that a pronoun bound by the possessor is as acceptable as a pronoun bound by a subject, but that both are only moderately acceptable.

We also investigated the influence of the pragmatic factors verb class, definiteness, and referentiality on coreference in PNPs. The underlying theoretical assumption is that coreference for exempt anaphors is governed by pragmatic factors, rather than by structural constraints. In binding configurations (i) and (ii), we found a significant effect of verb class: the acceptability of pronouns was reduced for [+existence] accomplishment verbs (see Figure 4). This accords with intuitions reported in the literature (see Introduction, (4)). Furthermore, we found a significant effect of definiteness: pronouns are more acceptable with definite PNPs than with indefinite ones (see Figure 5). However, the verb class effect and the definiteness effect were weak and did not change the overall acceptability pattern, i.e., the preference for anaphors over pronouns.

In configurations (iii) and (iv), we found that the pragmatic factor referentiality has a significant effect on the acceptability of pronouns, which were less acceptable if the antecedent is a quantified NP, compared to cases where the antecedent is a name or a definite NP (see Figure 6). Again, this effect was weak and did not change the overall pattern, i.e., the fact that both pronouns and anaphors were highly acceptable in configurations (iii) and (iv). Finally, we failed to find any effect of referentiality in configurations (v) and (vi) (see Figure 7).

Conclusions

Experiment 1 was a control study that made a methodological contribution. The results showed that the experimental paradigm of magnitude estimation, previously only used for grammaticality judgment tasks, can be applied successfully to coreference judgments, which form the empirical basis of binding theory.

Building on this result, Experiment 2 used magnitude estimation to provide crucial data regarding binding in PNPs, which have been the subject of much research in the syntactic literature. The results provide an example of how experimentation can be used as a tool to settle debates in linguistic theory.

More specifically, Experiment 2 aimed to clarify the empirical status of exempt anaphors and provide data to distinguish between structural and pragmatic accounts of exempt anaphors. The results show that structural factors govern the binding possibilities in PNPs, while pragmatic factors play only a limited role. However, the structural factors identified differ from the ones standardly assumed. We found that (i) an anaphor can be bound from outside the PNP, even if there is a possessor in the PNP, (ii) anaphors and pronouns bound by the subject are equally acceptable when there is a possessor, (iii) pronouns are only moderately acceptable when there is no possessor, and (iv) pronouns bound by the possessor are also moderately acceptable.

Finding (i) is the most theoretically interesting one, and has recently been confirmed in an eye-tracking experiment (Runner, Sussman, & Tanenhaus, 2000). It falsifies a major prediction of all binding theories by showing that structural factors (subject/no subject, possessor/no possessor) fail to influence the binding of anaphors in PNPs. This means that the role of structural factors is even smaller than envisaged by proponents of pragmatic accounts. For pronouns, however, there is a structural effect, viz., they are more acceptably bound by the subject if there is a possessor NP.

In our view, the best way to understand this result is by making reference to the notion of predication (Pollard & Sag, 1994; Reinhart & Reuland, 1993). An anaphor must be bound by a dominating coargument of the predicate that selects for the anaphor, if there is one. For example, an anaphor that is in the object position of a matrix clause must be bound by the subject, because the subject position dominates the object position: both the subject and object are arguments of the same predicate, i.e., the predicate needs the subject and object to satisfy its syntactic and semantic requirements. But the possessor of a PNP is not an argument of the head, as the head does not require it (i.e., pictures do not *necessarily* have possessors). This observation correctly accounts for the full acceptability of anaphors in PNPs, with or without possessors, and the necessity for local binding when anaphors are in matrix argument positions (as in (1)).

We can also use the notion of predication to understand the pattern for [+existence] accomplishment verbs, as in (4) *Hanna$_i$ took a picture of *her$_i$/herself$_i$*, without positing a covert possessor. It is possible that speaker-hearers treat expressions like *take a picture* as one predicate, in which case the anaphor or pronoun in such examples is actually a coargument of the subject and governed by Principle A or B, respectively. Runner (to appear) argues for just such an analysis of predicates like *take a picture* based on syntactic and semantic evidence.

References

Bard, E. G., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, *72*(1), 32–68.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, N., & Lasnik, H. (1995). The theory of principles and parameters. In *The minimalist program* (pp. 13–127). Cambridge, MA: MIT Press.

Corley, M., & Scheepers, C. (in press). Syntactic priming in English: Evidence from response latencies. *Psychonomic Bulletin*.

Cowart, W. (1997). *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks, CA: Sage Publications.

Gordon, P. C., & Hendrick, R. (1997). Intuitive knowledge of linguistic coreference. *Cognition*, *62*, 325–370.

Keller, F., & Alexopoulou, T. (2001). Phonology competes with syntax: Experimental evidence for the interaction of word order and accent placement in the realization of information structure. *Cognition*, *79*(3), 301–372.

Keller, F., Corley, M., Corley, S., Konieczny, L., & Todirascu, A. (1998). *WebExp: A Java toolbox for web-based psychological experiments* (Technical Report No. HCRC/TR-99). University of Edinburgh: Human Communication Research Centre.

Kuno, S. (1987). *Functional syntax: Anaphora, discourse and empathy*. Chicago: University of Chicago Press.

Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, *24*(4), 657–720.

Runner, J. T. (to appear). When Minimalism isn't enough: An argument for argument structure. *Linguistic Inquiry*.

Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2000). *Binding reflexives and pronouns in real-time processing*. Poster at the 13th Annual CUNY Conference on Human Sentence Processing, San Diego, CA.

Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: John Wiley.