Constraints on Linguistic Coreference: An Experimental Investigation of Exempt Anaphors

Frank Keller
keller@cogsci.ed.ac.uk
School of Cognitive Science
Division of Informatics, University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW, UK

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

Abstract

Previous experimental studies (Gordon & Hendrick, 1997) demonstrated that native speakers' judgments of the coreference possibilities for (non-reflexive) pronouns systematically diverge from the predictions of linguistic binding theory. We extend these results to the coreference possibilities for syntactic anaphors (reflexives). In particular, we deal with the binding of anaphors and pronouns as they occur in picture NPs. We show, contrary to the predictions of binding theory, that in picture NPs, anaphors are systematically preferred over pronouns. Furthermore, such anaphors can take remote binders, even when there is a local, intervening nominal. This entails that such anaphors are always exempt from binding theory, contrary to the claims in the literature. We also demonstrate that the same factors that influence the exempt status of anaphors play a role in extraction from picture NPs. This result indicates that both phenomena should be mediated by the same grammatical representations.

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) allows one 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 problem in linguistic theory, viz., anaphora that seem to be exempt from binding theory. The theoretical literature on this phenomenon is characterized by considerable data disagreements, i.e., disagreements about the relevant linguistic intuitions. We will demonstrate how the use of experimentally elicited intuitions can help to resolve such data disputes and to discover interesting new facts for linguistic theory.

Binding Theory

Binding theory is the module of grammar that regulates the interpretation of noun phrases (NPs). Three types of noun phrases are generally distinguished: (a) full NPs such as *Hanna* or *the woman*, (b) pronouns such as *he* and *her*, and (c) anaphors, i.e., reflexives such as *herself*, or reciprocals such as *each other*.

The task of binding theory is to determine which noun phrase can be *coreferential*, i.e., refer to the same individual. Coreference is normally indicated with subscripts:

(1) a. Hanna$_i$ admires *her$_i$/herself.
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 situation for the reflexive is exactly opposite; *Hanna* and *herself* can be coreferential in (1a), but not in (1b).

There are structural conditions that determine the binding possibilities of anaphors and pronouns. Principle A of binding theory captures the binding requirements for anaphors; it states that an anaphor has to be bound within a certain local domain (Chomsky, 1986). Principle B, on the other hand, states that pronouns cannot be bound within the same local domain. It follows that anaphors and pronouns are in complementary distribution, i.e., anaphors can be bound when pronouns cannot be bound, and vice versa.

Exempt Anaphors

It has been observed by a number of authors (e.g., Pollard & Sag, 1994; Reinhart & Reuland, 1993) that in certain configurations, anaphors are exempt from binding theory. In such cases, the anaphor is not subject to Principle A. Relevant configurations include picture NPs (PNPs) without possessors, as illustrated in (2a), where the binding of an anaphor and a pronoun are both acceptable. When there is a possessor in the PNP, the relevant domain for anaphoric binding is the NP, and anaphors are claimed to be ungrammatical in sentences like (2b), while pronouns are fine. When there is no possessor in the PNP, both pronouns and anaphors are claimed to be acceptable.

(2) a. Hanna$_i$ found a picture of her$_i$/herself$_i$.
    b. Hanna$_i$ found Peter's picture of her$_i$/*herself$_i$.

Based on such data, authors like Pollard and Sag (1994) have argued that Principle A should be formulated so as not to apply to anaphors in sentences such as (2a). The assumption is that the binding properties of such anaphors are governed by non-syntactic factors, including processing and discourse constraints.

The present study has a triple purpose. First, we attempt to clarify the empirical status of exempt anaphors. By conducting a study with linguistically naive native speakers we can determine whether anaphors and pronouns are perceived as equally acceptable in relevant configurations, like the one in (2a). This question will be addressed in Experiment 1. (For other studies demonstrating the usefulness of experimental data in clarifying binding facts, see Cowart, 1997; Gordon & Hendrick, 1997, 1998.)

The second purpose is to shed light on the factors that influence the distribution of pronouns and exempt anaphors. Such factors include the referentiality of the binder, the definiteness of the PNP, and the aspectual class of the matrix verbs. These factors will also be investigated in Experiment 1.

Our third aim is to determine if the factors that influence the exempt status of an anaphor also lead to systematic acceptability differences in other syntactic phenomena. We investigate extraction from PNPs, a phenomenon that has traditionally been treated separately from binding. Our hypothesis is that the same set of factors influence acceptability in both constructions. If correct, this hypothesis would entail that binding and extraction should receive a unified linguistic account. Extraction will be the topic of Experiment 2.

Experiment 1: Binding in Picture NPs

This experiment has two subdesigns. The first one investigates how the exempt status of an anaphor is influenced by the definiteness of the PNP and by the aspectual class of the matrix verb. As an example of definiteness consider the minimal pair in (3): the PNP in (3a) is indefinite and the one in (3b) is definite.

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

The factor verb class is illustrated in example (4): *find* and *lose* are examples of achievement verbs, while *take* and *destroy* are accomplishment verbs; *find* and *take* are [+existence], while *lose* and *destroy* are [−existence].

The second subexperiment was designed to test the influence of an intervening NP, as illustrated by the minimal pair in (5). The intervention of a potential binder was identified by both Asudeh (1998) and Pollard and Sag (1994) as a relevant factor in determining the exempt status of an anaphor. According to Pollard and Sag (1994), the anaphor in (5a) is exempt because it does not have a potential referential binder in its local domain (the PNP), whereas the anaphor in (5b) is not exempt since the PNP contains a local referential nominal. The second subexperiment also tested the influence of the referentiality of the binder, as illustrated in (6). We also included a control condition where the intervening NP is the binder, as shown in (7):

(5) a. Hanna$_i$ found a picture of her$_i$/herself$_i$.
    b. Hanna$_i$ found Peter's picture of her$_i$/herself$_i$.
(6) 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$_i$.
(7) a. Hanna$_i$ found Peter's picture of her$_i$/herself$_i$.
    b. Hanna found Peter's$_i$ picture of him$_i$/himself$_i$.

In the present experiment, we elicited acceptability judgments for both the anaphor and the pronoun in configurations like the ones in (3)–(7). Our aim is to test if the factors definiteness, verb class, referentiality, and the intervention of a binder have a significant influence on the binding theoretic status of a given configuration.

Predictions

Based on the binding literature, we predict that an anaphor and a pronoun are equally acceptable in examples like (2a). This means that we should fail to find a main effect of NP type (anaphor or pronoun). We also expect that the intervention of a potential binder (see (5)) influences the exempt status of an anaphor, based on the theoretical claims by Asudeh (1998) and Pollard and Sag (1994). Hence we should find a significant interaction of intervention and NP type.

While previous experimental studies showed that referentiality can affect binding (Gordon & Hendrick, 1998), there is no previous experimental work dealing specifically with exempt anaphors or with factors such as definiteness and verb class. However, there is some discussion of such effects in the theoretical literature (Chomsky, 1986; Kuno, 1987; Pollard & Sag, 1994; Reinhart & Reuland, 1993), based on which we predict referentiality, definiteness, and verb class to influence binding in PNPs. This means that our experiment should show interactions between NP type and these three factors.

Finally, Principle A predicts that anaphors lose their exempt status in the control condition (see (7)), where there is a referential potential local binder inside the PNP. For the indicated coreference, binding theory predicts that (7a) should be ungrammatical with the anaphor and grammatical with the pronoun, while (7b) is grammatical with the anaphor and ungrammatical with the pronoun. This should manifest itself in the experiment as an interaction of binder and NP type.

Method

Subjects Fifty-two native speakers of English participated in the experiment. All participants were naive to syntactic theory.

Materials The experimental materials included two subdesigns. The first subdesign used the factors definiteness (DEF), verb class (VERB), and anaphor (ANA). There were two levels for DEF (definite, indefinite, see (3)), three levels for VERB (achievement [+existence], accomplishment [+existence], accomplishment [−existence] (see (4a), (4c), (4d)), and two levels for ANA (NP type, i.e., anaphor or pronoun). This yielded a total of DEF × VERB × ANA = $3 \times 2 \times 2 = 12$ cells.

The second subdesign included the factors referentiality (REF), binder (BIND), and NP type (ANA). There were three levels for REF (proper name, definite NP, quantified NP, see (6)), two levels for BIND (remote or local binder, see (7)), and two levels for ANA (anaphor, pronoun), yielding a total of REF × BIND × ANA = $3 \times 2 \times 2 = 12$ cells. Four lexicalizations were used for each of the cells, which resulted in a

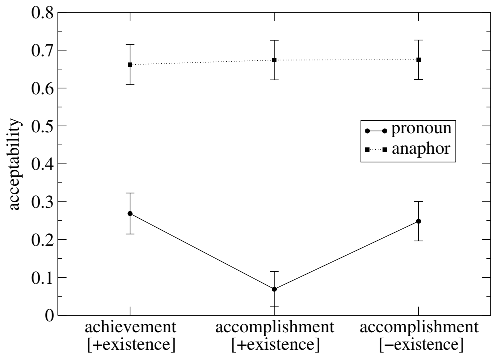

Figure 1: Interaction of VERB and ANA

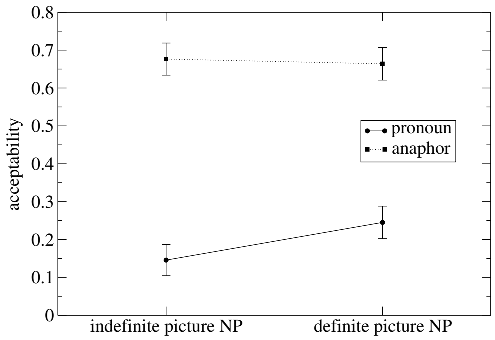

Figure 2: Interaction of DEF and ANA

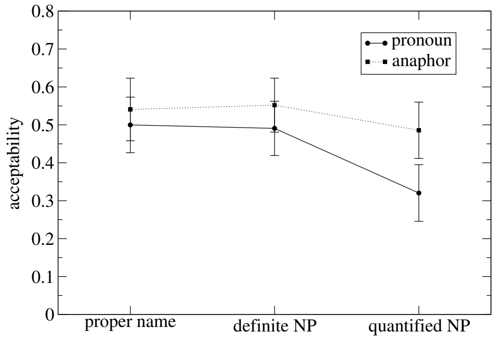

Figure 3: Interaction of REF and ANA (remote binder)

total of 96 stimuli. A set of 24 fillers was used, designed to cover the whole acceptability range.

Procedure The method used was magnitude estimation (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 judge the acceptability of coreference. Following Gordon and Hendrick (1997), this 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.

The experiment started with a training phase designed to familiarize subjects with the ME task. Subjects had to estimate 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 for linguistic stimuli. Finally, subjects had had to judge the experimental items. The 96 stimuli were placed in a Latin square design, generating four lexicalizations at 24 items each. Each subject judged one lexicalizations and 24 fillers, i.e., a total of 48 items. Subjects were assigned to lexicalizations at random, and stimulus presentation was randomized for each subject.

Results

The data were normalized by dividing each numerical judgment by the modulus value that the subject had assigned to the reference sentence. This operation creates a common scale for all subjects. All analyses were carried out on the geometric means of the normalized judgments, as is standard for ME data (Bard et al., 1996; Cowart, 1997).

The ANOVA on the first subexperiment yielded a main effect of VERB (verb class) ($F_1(2,102) = 9.345$, $p < .0005$; $F_2(2,6) = 4.839$, $p = .056$): [+existence] accomplishment verbs like *take* were significantly less acceptable (mean = .3715) than [−existence] accomplishment verbs like *destroy* (mean = .4653) or [+existence] achievement verbs like *find* (mean = .4616). The main effect of DEF (definiteness) was small and only significant by subjects ($F_1(1,51) = 7.927$, $p = .007$; $F_2(1,3) = 1.207$, $p = .352$). Definite PNPs (mean = .4546) were more acceptable than indefinite ones (mean = .4110). We also found a large and highly significant main effect of ANA (NP type) ($F_1(1,51) = 137.471$, $p < .0005$; $F_2(1,3) = 105.005$, $p = .002$). Anaphors (mean = .6702) were more acceptable than pronouns (mean = .1954).

The ANOVA 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 1, which shows that there is 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 ($F_1(1,51) = 11.849$, $p = .001$; $F_2(1,3) = 2.168$, $p = 0.237$). Figure 2 shows that the acceptability for pronouns is increased for definite PNPs. The interaction of VERB and DEF, as well as the three-way interaction of VERB, DEF, and ANA failed to be significant.

The ANOVA on the second subexperiment revealed a main effect of BIND (remote or local binder), which however was significant by subjects only ($F_1(1,51) = 7.851$, $p = .005$; $F_2(1,3) = 4.284$, $p = 0.130$). A remote binder (mean = .4816) was more acceptable than a local binder (mean = .4085). The factor REF (referentiality) was highly significant ($F_1(2,102) = 68.244$, $p = .001$; $F_2(2,6) = 12.197$, $p = .008$); quantified binders like *each woman* (mean = .4008) were less acceptable than non-quantified binders such as *Hanna* (mean = .4672) or *the woman* (mean = .4670). Finally, we replicated the effect of ANA found in the first subexperiment ($F_1(1,51) = 68.244$, $p < .0005$; $F_2(1,3) = 45.725$, $p = .007$). Again, anaphors (mean = .5800) were more acceptable than

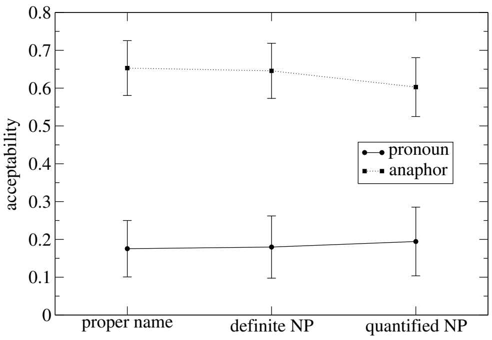

Figure 4: Interaction of REF and ANA (local binder)

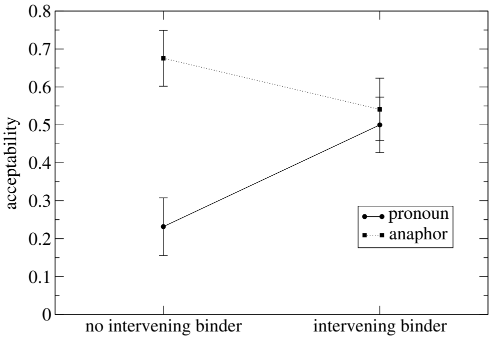

Figure 5: Interaction of INT and ANA

pronouns (mean = .3101).

The ANOVA also demonstrated a significant interaction of BIND and REF ($F_1(2,102) = 3.966$, $p = .022$; $F_2(2,6) = 10.638$, $p = .011$). The interaction of BIND and ANA was significant by subjects and marginal by items ($F_1(1,51) = 35.051$, $p < .0005$; $F_2(1,3) = 6.274$, $p = .087$). Finally, a three-way interaction of BIND, REF, and ANA was also obtained ($F_1(2,102) = 4.041$, $p = .020$; $F_2(2,6) = 4.543$, $p = .063$). This interaction is graphed in Figures 3 and 4. An inspection of Figure 3 shows that in the remote binder condition, pronouns and anaphors are equally acceptable if the binder is a proper name or a definite NP. However, if the binder is a quantified NP, the acceptability for pronouns decreases. There is no such effect in the control condition (local binder, see Figure 4). A post-hoc Tukey test on the BIND/REF/ANA interaction confirms this observation: for the remote binder condition, the difference between pronoun and anaphor is not significant for proper names and definite NPs, but reaches significance for the quantified NPs (by subjects only, $\alpha < .01$). For the local binder condition, on the other hand, the difference between pronoun and anaphor is significant for all three binders (by subjects and by items, $\alpha < .05$).

A comparison of Figures 1 and 2 and Figures 3 and 4 shows that PNPs only have exempt status if there is an intervening potential binder. If there is no intervening binder, pronouns are highly unacceptable with PNPs—we get essentially the same acceptability pattern as in the case of a local binder. To confirm this observation, we conducted an ANOVA on the data that overlapped from the two subexperiments (see (5) for an example). The factors were ANA (pronoun or anaphor) and INT (intervening potential binder or not). A main effect of INT was found, which however was significant by subjects only ($F_1(1,51) = 5.142$, $p = .028$; $F_2(1,3) = 1.747$, $p = 0.278$). We also found a main effect of ANA, which was significant by subjects and marginal by items ($F_1(1,51) = 33.181$, $p < .0005$; $F_2(1,3) = 6.987$, $p = 0.077$). Crucially, there was a significant interaction of INT and ANA ($F_1(1,51) = 35.432$, $p < .0005$; $F_2(1,3) = 15.608$, $p = 0.029$). This interaction is graphed in Figure 5.

Discussion

This experiment demonstrated that binding in PNPs is not equally acceptable for anaphors and pronouns. For examples such as (2a), we found a main effect of ANA (NP type), which shows that pronouns are consistently less acceptable than anaphors. Binding theory, as commonly formalized in various frameworks, expects pronouns to be grammatical in PNPs, but has to take some extra measures to account for the grammaticality of anaphors in the same configurations. For example, Chomsky (1986) introduces the notion of counterfactual coindexation to extend the domain of anaphoric binding in such cases. Pollard and Sag (1994) exempt anaphors in PNPs from binding theory altogether (so long as there is no referential possessor in the PNP). They argue that the reference of such anaphors is governed by discourse and processing constraints, which they never explicitly spell out (although they do give a sketch of certain relevant factors). Our results suggest that anaphors should actually be treated as the base case and that it is pronouns that are marginal and exceptional in PNPs.[1]

We also tested cases where another potential binder intervenes between the pronoun or anaphor and its antecedent, such as in (5). In this case, the acceptability of pronouns and anaphors is not significantly different. We found that this finding only holds if the binder is referential, i.e., a proper name or a definite NPs. There is a reduction in the acceptability of the pronoun if the binder is a quantified NP. This was evidenced by the interaction of REF and ANA in our data.

Crucially, we found an interaction of INT and ANA (see Figure 5). Our results show that there is no significant difference in the acceptability of anaphors and pronouns in such sentences. Again this contradicts claims in the theoretical literature. When there is an intervening binder, as in (7) above, binding theory predicts that only a pronoun should be able to have an antecedent outside the PNP (as in (7b)). Although the anaphor decreases in acceptability and the pronoun increases in acceptability compared to the case with no intervening binder, they are equally acceptable and it is simply not true that anaphors are ungrammatical in such cases, if pronouns are to be considered grammatical.

We also found an interaction of DEF and ANA and an interaction of VERB and ANA. This demonstrated that the acceptability of pronouns improves if the PNP is definite or if the matrix verb is a [+existence] achievement verb or a [−existence] accomplishment verb. However, this improvement in acceptability does not compensate for the general unacceptability of pronoun binding in PNPs (see Figures 1 and 2). We found that unacceptability of pronouns in a PNP increases if the PNP is complement to a [+existence] accomplishment verb, in line with the claims in theoretical literature

[1] One could ask whether our results could be an artifact of the fact that we used linguistically naive speakers, which failed to apply the concept of coreference as intended. Note that the filler we ran in our experiment were a replication of Gordon and Hendrick's (1997) Experiments 1–4, which tested very basic binding facts (such as the ones in (1)). The results we obtained closely matched Gordon and Hendrick's original results, which indicates that our subjects did use the concept of coreference correctly.

(Chomsky, 1986; Chomsky & Lasnik, 1995; Reinhart & Reuland, 1993).

In our control condition, where the binder was inside the NP (see (7b), we found that anaphors are highly acceptable, while pronouns are highly unacceptable, as predicted by binding theory. Here, we failed to find an effect of referentiality, i.e., referential and quantified NPs were equally unacceptable.

In the following experiment, we investigate the influence of the factors DEF, REF, and VERB on a different construction involving PNPs, viz., extraction. Our hypothesis is that a common set of factors influences both binding and extraction in PNPs, thus indicating that both phenomena are mediated by the same grammatical mechanism.

Experiment 2: Extraction from Picture NPs

Extraction from PNPs has attracted considerable interest in the theoretical literature (Kluender, 1992; Erteschik-Shir, 1981; Fiengo, 1987), but has largely been treated separately from binding in PNPs. Insofar as *wh*-extraction also involves a kind of referential dependency (the extracted *wh*-word is questioning an aspect of the complement in the PNP), we expect that binding and extraction are affected by the same factors, although not necessarily in the same way.

Complex NPs are standardly assumed to be islands for extraction. PNPs, however, constitute well-known counterexamples to this assumption, as they allow island violations in certain cases. A number of factors are know to influence the island status of PNPs. For instance, Kluender (1992) and Fiengo (1987) observe that definiteness has an influence on extractability: extraction from indefinite PNPs is more acceptable than extraction from definite ones (see (8)). This mirrors the definiteness effect for pronouns found in Experiment 1. In that experiment, binding of pronouns in indefinite PNPs was found to be *less* acceptable than binding of pronouns in definite PNPs.

(8) a. Which friend has Thomas painted a picture of?
b. Which friend has Thomas painted the picture of?

Extractability also depends on the semantic class of the matrix verb. Aspectual information seems to be a main factor here: state verbs are more acceptable than activity verb (see (9)), while for accomplishments and achievements, a [+existence] verb is more acceptable than a [−existence] verb (see (10) and (11)). Again, these observations are not unexpected given the effect of verb on binding; but again in Experiment 1 pronouns behaved in the opposite fashion: pronominal binding is *worse* in [+existence] accomplishment verbs.

(9) a. Which friend has Thomas owned a picture of?
b. Which friend has Thomas analyzed a picture of?
(10) a. Which friend has Thomas painted a picture of?
b. Which enemy has Thomas torn up a picture of?
(11) a. Which friend has Thomas found a picture of?
b. Which friend has Thomas lost a picture of?

A third factor influencing the acceptability of extraction from PNPs is the referentiality of the extracted NP. It has been claimed (Kluender, 1992) that referential NPs like *which friend* are more extractable than non-referential ones like *how many friends*:

(12) a. Which friend has Thomas painted a picture of?
b. How many friends has Thomas painted a picture of?

This effect is not unlike the referentiality effect observed for binding in examples like (6). Quantified NPs like *each woman* share with *wh*-phrases like *how many friends* the property of being non-referential. On the other hand, *wh*-phrases like *which friend* are referential, as are the definite NPs or proper names examined as binders in Experiment 1.

Predictions

Based on claims from the relevant theoretical literature (Kluender, 1992; Erteschik-Shir, 1981; Fiengo, 1987) and on our hypothesis that the same set of factors influence both binding and extraction in PNPs, we expect that all of the factors included in the present experiment (DEF, VERB, and REF) should have a significant effect on the acceptabilty extraction from PNPs.

Method

Subjects Fifty-one subjects from the same population as in Experiment 1 participated. None of them had previously taken part in Experiment 1.

Materials Training and Practice Materials were designed in the same way as in Experiment 1.

The experiment crossed the factors definiteness (DEF), referentiality (REF), and verb class (VERB). There were two levels for DEF (definite, indefinite, see (8)), two levels for VERB (accomplishment [+existence], accomplishment [−existence], see (10)), and two levels for REF (referential, non-referential, see (12)). This yielded a total of DEF × REF × VERB = 2 × 2 × 2 = 8 cells.

Procedure The method used was ME of linguistic acceptability, with the same experimental protocol as in Experiment 1. A modified version of the instructions of Experiment 1 was used. Subject were now asked to judge acceptability instead of coreference.

The 32 experimental stimuli were placed in a Latin square design, generating four lexicalizations at eight items each. Each subject judged one lexicalizations and 24 fillers, i.e., a total of 32 items. Subjects were assigned to lexicalizations at random, and stimulus presentation was randomized for each subject.

Results

The data were normalized as in Experiment 1. An ANOVA showed that the factor DEF was significant ($F_1(1,50) = 13.005$, $p = .001$; $F_2(1,3) = 34.406$, $p = .010$): extraction from indefinite PNPs (mean = .0335) was more acceptable than extraction from definite ones (mean = −.0084). A main effect of REF was also found ($F_1(1,50) = 29.338$, $p < .0005$; $F_2(1,3) = 31.799$, $p = .011$): extraction of referential *wh*-phrases (mean = .0460) was more acceptable than extraction of non-referential ones (mean = −.0208). Finally, there was a main effect of VERB ($F_1(1,50) = 29.549$, $p < .0005$; $F_2(1,3) = 108.546$, $p = .002$): verbs of the type [+existence] (mean = .0486) were more acceptable than [−existence] verbs (mean = −.0234). All interactions failed to be significant.

Discussion

Based on the results of Experiment 1, the present experiment investigated extraction from PNPs. Based on our assumption that the same grammatical representation underlies binding and extraction, we predicted these same factors that influenced the acceptability of in PNPs should also surface in the present experiment.

As predicted, we found that the factors definiteness, verb class, and referentiality have a significant influence on the acceptability of extraction from PNPs. On the one hand, this result provides an experimental confirmation of the claims regarding these factors in the theoretical literature, which typically rely on an intuitive data base. On the other hand, the result supports our initial hypothesis that the same underlying representation should account for both binding and extraction. We will discuss this in more detail in the next section.

Anaphors, Pronouns, and Traces

Taken together, the results of Experiments 1 and 2 support the hypothesis that the same set of factors influence both binding and extraction in PNPs. In effect there are three patterns. With respect to the factors DEF and VERB, extraction and the binding of pronouns exhibit opposing patterns, while the binding of anaphors is insensitive to these factors, and therefore exhibits a third pattern. This provides some motivation for the categorical distinction made between pronominals, anaphors and traces in Chomsky (1982).

However, it does not commit us to accepting syntactic traces as representational devices; a more restrictive hypothesis is that the target of extraction has to be marked in some manner so that we can make generalizations about it regarding the factors we have isolated here. Such an analysis is put forward by Bouma, Malouf, and Sag (2000) in the framework of Head-Driven Phrase Structure Grammar (HPSG; Pollard & Sag, 1994). This account is compatible with our results as it treats extraction as a discrepancy between the argument structure of a given lexeme (a verb for example) and its syntactic combinatorial requirements and deals with extracted elements as a special syntactico-semantic object. Bouma et al.'s (2000) approach does not treat the site of extraction as a terminal node that is satisfying the combinatorial requirements of the verb by projecting a phrase that the verb subcategorizes for; thus Bouma et al. do not postulate traces for extraction.

If we accept this analysis of extraction in HPSG, we have a theory that represents both binding and extraction at the level of argument structure. Our experiments have shown that these phenomena are sensitive to similar factors. We take this as evidence that it is desirable to deal with both phenomena using the same grammatical representations, as is the case in HPSG.

Conclusions

We presented the results of two experimental studies that used magnitude estimation to investigate binding and extraction in picture NPs. The results showed that in picture NPs, anaphors are systematically preferred over pronouns and that such anaphors can take remote binders, even when there is a local, intervening nominal. This entails that such anaphors are always exempt from binding theory. We argued that these findings are unexpected under typical formulations of binding theory, and provide evidence for extending the set of configurations where anaphors are exempt from binding principles. Such a move can build on existing proposals regarding exempt anaphors advanced by Asudeh (1998) and Pollard and Sag (1994).

The second experimental finding was that the same factors that influence the exempt status of anaphors play a role in extraction from picture NPs. This result provides some motivation for the categorical distinction made between pronominals, anaphors and traces in Chomsky (1982). However, we argued that the shared behavior of binding and extraction can be better captured in a framework that relies on a separate level of argument structure to handle both binding and extraction (Bouma et al., 2000).

On a more general level, we have demonstrated that experimentally elicited acceptability judgments can clarify binding facts, which is an example of how experimentation can help to uncover interesting data for linguistic theory.

References

Asudeh, A. (1998). *Anaphora and argument structure: Topics in the syntax and semantics of reflexives and reciprocals.* Unpublished master's thesis, University of Edinburgh.

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

Bouma, G., Malouf, R., & Sag, I. A. (2000). Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory*. (to appear)

Chomsky, N. (1982). *Some concepts and consequences of the theory of Government and Binding.* Cambridge, MA: MIT Press.

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.

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

Erteschik-Shir, N. (1981). On extraction from noun phrases (picture noun phrases). In A. Belletti, L. Brandi, & L. Rizzi (Eds.), *Theory of markedness in generative grammar: Proceedings of the 4th GLOW Conference* (pp. 147–169). Scuole Normale Superiore di Pisa.

Fiengo, R. (1987). Definiteness, specificity, and familiarity. *Linguistic Inquiry*, *18*, 163–166.

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

Gordon, P. C., & Hendrick, R. (1998). Nondefinite NP anaphora: A reappraisal. In M. C. Gruber, D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *Papers from the 34th Meeting of the Chicago Linguistic Society* (Vol. 1: The Main Session, pp. 195–210). Chicago.

Kluender, R. (1992). Deriving island constraints from principles of predication. In H. Goodluck & M. Rochemont (Eds.), *Island constraints: Theory, acquisition and processing* (pp. 223–258). Dordrecht: Kluwer.

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.

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

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