

Running head: REPRESENTING REFERENTS OF PLURAL EXPRESSIONS

2002 LANGUAGE AND COGNITIVE PROCESSES, 17, 405-450

Representing Referents of Plural Expressions and Resolving Plural Anaphors

Barbara Kaup Stephanie Kelter,

Hamburg University Technical University of Berlin, Germany

Christopher Habel,

Hamburg University, Germany

Abstract

In this paper, two views for mentally representing referents of plural expressions are contrasted, (a) the atomic-tokens view, according to which the denotation of a plural expression is represented by a number of distinct tokens, and (b) the assemblage-token view, according to which it is represented as a single whole. Linguistic data suggest that the atomic-tokens view is appropriate for partitive plural NPs (e.g., most of the orphans; both of the cars) and functionally similar expressions (e.g., most orphans; both), whereas the assemblage-token view is appropriate for unmarked plural expressions (e.g., the orphans, my cars, they). This conjecture was investigated in 2 experiments with German participants, contrasting the pronoun sie (they) and the partitive pronoun beide (both). In Experiment 1, off-line tasks were used to investigate whether the interpretation of a sentence's predicate depends on the pronoun used in the grammatical subject position. In Experiment 2, participants read texts containing a pronoun in different contexts, and sentence reading times were measured to find out whether or not the pronoun triggered a grouping of the individuals referred to. The results support the hypotheses derived from linguistic analyses.

Representing Referents of Plural Expressions and Resolving Plural Anaphors

One of the most important functions of language is to exchange information about states of affairs. Crucial for this enterprise are referential expressions. Referential expressions are the means by which the speaker specifies the entities about which he or she wants to provide information, and by means of which the listener identifies these entities. In most philosophical and logico-semantic theories of natural language, reference is regarded as a relationship between a linguistic expression and an entity in the real or a possible world. From a cognitive-science point of view, however, this relationship is indirect at best. This view suggests that it is the human mind that establishes the connection between linguistic and non-linguistic entities. Thus, cognitive research analyzes linguistic reference in terms of a relationship between mental representations of linguistic expressions and mental representations of non-linguistic entities. Hence, despite considerable differences in their basic assumptions, virtually all modern theories of discourse comprehension postulate a non-linguistic level of representation, the elements of which are tokens standing for the referents of linguistic expressions. These non-linguistic representations are dubbed differently in the various theories, for example: mental models (Garnham, 1996; Glenberg & Langston, 1992; Johnson-Laird, 1983), situation models (van Dijk & Kintsch, 1983; Morrow, 1994; Zwaan & Radvansky, 1998), explicit focus (Garrod & Sanford, 1998), or mental structures (Gernsbacher, 1990, 1997).

A particular difficulty that a cognitive approach to reference is confronted with is that the elements that enter into the relationship to be investigated are themselves not well known and need to be explored. This holds not only for the linguistic representations but for the non-linguistic representations as well. The human mind does not merely represent entities that are already given in the external world, but rather plays an active role and imposes a structure onto the external world. In other words, the human mind construes the entities and projects them onto the world (Jackendoff, 1983, chap. 2, 1987, chap. 7; see also Johnson-Laird, 1983, chap. 9).

Even more, the structure imposed is probably not fixed. Depending on the particular situation, the same stimuli may be conceived of as different kinds of entities. For instance, walking through a wooded area, we can conceive of what we see as a number of individual trees or as a forest (cf. Bloom, 1996). Hence, cognitive research on linguistic reference has to investigate the cognitive ontology drawn on in discourse and, more specifically, the factors that determine what kinds of entities are differentiated in the mind in a particular situation.

A tremendous number of experimental studies has been devoted to specifying the dimensions of states of affairs represented at the non-linguistic level during text comprehension (for a review see Zwaan & Radvansky, 1998). Yet, comparatively little attention has been directed to the question of what kinds of entities are differentiated in these representations. Most empirical studies have been concerned with singular expressions referring to discrete physical objects (e.g., humans, animals, and household items). This is understandable in that discrete physical objects seem to play a prominent role in cognition, as evidenced clearly in research on language acquisition (Bloom & Kelemen, 1995; Markman, 1990; Shipley & Shepperson, 1990). However, it is an open question to what extent the findings can be generalized to other types of referential expressions. It is conceivable that places, events, substances, and abstract entities are mentally construed in different ways (cf. Medin, Lynch, & Solomon, 2000), so that from a functional point of view different kinds of tokens have to be distinguished. Moreover, when considering singular expressions with a generic reading (e.g., the whale in the whale is a mammal but lives in the water) and collective terms (e.g., the football team) it even appears questionable whether there is always a one-to-one mapping between tokens and individual entities (Gelman & Tardiff, 1998; Landman, 1989).

The present paper is concerned with plural expressions such as the houses, most orphans, or they. At first glance it seems clear that for a plural expression the mental representation at the non-linguistic level is comprised of several tokens. However, analyzing plural expressions in

greater detail reveals that matters are not that simple. The idea that the referent is mentally represented by several tokens holds for certain types of plural expressions only. For other types of plural expressions, an alternative view, according to which the individuals are mentally aggregated, is more adequate.

There are already some experimental studies on the processing of plural expressions. Instead of starting with a review of these studies, however, we shall first lay out the issue of plural reference more systematically, mainly by drawing upon linguistic observations and arguments. In the second part of the paper, the resolution of plural anaphors will be discussed in more detail. In this context, the available experimental findings will be reported as well as the results of two new experiments.

In order to avoid tackling too many new issues at once, we will limit considerations to referentially used plural expressions denoting concrete things. Furthermore, since our interest is in reference, we will focus on referentially used plural noun phrases (NPs) such as the men and the lawyers in (1) and disregard indefinite NPs occurring in predicates such as lawyers in (2).

(1) The men visited the lawyers.

(2) The men are lawyers.

Representing the Referents of Plural Expressions

Imagine that Dr. Larch, telling his co-workers about his trip with ten orphans to the circus this afternoon, says

(3) The orphans laughed.

What is the structure of the non-linguistic representation constructed for this utterance? One possibility is that there is a distinct token for each of the orphans and to each token the predicate 'laughing' is assigned (see Figure 1). We will call this view the atomic-tokens view. It is essentially what Johnson-Laird proposed (1983, pp. 442). It also seems to correspond to the implicit assumptions on "notional plurality" made in psycholinguistic studies of number

agreement (e.g., Bock, Nicol, & Cutting, 1999; Vigliocco, Butterworth, & Garrett, 1996). Note that it cannot be assumed that the number of tokens always corresponds to the number of individuals denoted. First, a plural term may denote an arbitrarily large number of individuals. It is unlikely, however, that with a sentence such as (4) a distinct token is created for each orphan.

(4) Eighty-two orphans have grown up in the orphanage since its founding.

Johnson-Laird (1983) proposed that with a large number of individuals "a manipulable number of tokens is taken to represent the set of (...) individuals" (p. 443), possibly attaching a propositional-like label to the representation that specifies the precise number of individuals. Second, the number of individuals denoted by a plural term may be indeterminate. Uttering sentence (3), for instance, would have been perfectly acceptable even without having specified the number of orphans beforehand. It is conceivable that in cases such as these the listener sets up the number of tokens which, according to world knowledge, is most likely for the kind of individuals and the type of situation referred to. Studies investigating the interpretation of vague terms (e.g., few, many, a pile) suggest that people make use of their knowledge about the typical number of objects of a particular kind in a particular situation, if the actual number is not explicitly given (Hörmann, 1983; see also Clark, 1991).

In conclusion, according to the atomic-tokens view, the individuals denoted by a plural NP are represented by a number of tokens with each token representing one member of the set. Not in all cases does the number of tokens correspond to the actual number of individuals denoted, but even in these cases each token stands for just one member of the set.

From a theoretical perspective, the atomic-tokens view seems attractive because it needs to postulate only one kind of token, which moreover corresponds to the standard notion of a mental token (i.e., one token represents one discrete physical individual). However, this representational framework is certainly not the only possibility when dealing with plural expressions. It is conceivable that the intended referents of plural NPs are mentally construed at

a level of granularity that is coarser than the one for individual things. Thus, in an alternative approach, it could be assumed that a plural NP is used for and understood as denoting individuals that are conceived of as a single entity and are therefore represented by a single token (see Figure 2). We will refer to this type of token as an "assemblage token."

The two accounts do not only differ in their assumptions as to the representation of the entities denoted by plural NPs but also with respect to the representation of the predicate expressed by the verb phrase of a sentence containing a plural subject NP (compare Figures 1 and 2). In contrast to the atomic-tokens view, the assemblage-token view implies that the predicate is conceived of as applying to an assemblage rather than as applying to atomic individuals. Therefore, let us take a closer look at the verb phrase (VP) of sentences with a plural subject NP.

Predicates ascribed to referents of plural expressions

In formal semantics, most theories of plurals assume that the discourse domain not only contains individuals in the usual sense (comparable to our atomic tokens), but also complex individuals (comparable to our assemblage tokens). These complex individuals are understood as mereological sums (e.g., Barker, 1992; Eschenbach, Habel, Herweg, & Rehkämper, 1989; Link, 1983; 1991; Massey, 1976) or special sets of atomic individuals (e.g., Kamp & Reyle, 1993; Landman, 1989; Winter, 2000). One of the main arguments for postulating these complex individuals are collective predicates.

Collective predicates are predicates that only apply to more than one individual. Prototypical examples are single joint actions of several individuals (e.g., The orphans assembled at the bus station, The dirty dishes piled up in the kitchen), or (non)correspondencies among the individuals (e.g., Today the orphans were dressed uniformly). Collective predicates indeed pose a problem for the atomic-tokens view. According to this view, there are only atomic tokens, none of which alone can take on the collective predicate. A proponent of the atomic-

tokens view could argue that collective predicates are of a very special type and may actually call for a different representational account than sentences with non-collective predicates. This would imply that the structure of the non-linguistic representation for a definite plural NP like the children is assumed to depend on the particular predicate of the sentence. If it is a non-collective predicate, the denotation of the plural subject NP is represented by a number of atomic tokens. If it is a collective predicate, the denotation is represented differently, possibly by an assemblage token. In fact, this idea comes close to what is proposed in many formal semantic theories, which distinguish between a distributive reading of sentences (the predicate is understood as applying to each atomic individual) and a collective reading (the predicate is understood as applying to the complex individual as a whole) (e.g., Kamp & Reyle, 1993; Link, 1991; Landman, 1989). However, a closer look at VPs involving internal arguments (object-NPs or NPs in prepositional phrases) reveals that the atomic-tokens view also encounters problems with sentences involving predicates that are usually considered distributive predicates¹.

Imagine Dr. Larch, continuing his report about the trip to the circus, has just told his audience that there was a fantastic act with five acrobats, and now wants to elaborate by describing the beginning of the act. The act started with the five acrobats riding into the ring, each sitting on a white horse. According to the atomic-tokens view, this situation can be described by denoting the acrobats through a plural NP and naming the predicate that applies to each of the acrobats. Hence, a sentence such as (5) should be appropriate.

(5) The acrobats rode into the ring on a white horse.

However, if Dr. Larch would utter (5), this would not be understood in the way intended, but rather as describing a situation in which the five acrobats are sitting together on the same white horse. Being a competent and cooperative speaker, Dr. Larch is actually more likely to say something like (6).

(6) The acrobats rode into the ring on white horses.

Yet, the atomic-tokens view cannot explain why (6) would be an appropriate description of the states of affairs. This view implies that upon hearing (6) the listener would definitely not represent the intended state of affairs (i.e., a state of affairs in which each acrobat is riding only one horse), since the predicate, which is taken to apply to each individual, is "riding white horses." Note that there is no morphosyntactic rule that calls for a plural object NP, when the grammatical subject is a plural NP; sentence (5) is grammatically well-formed, but simply does not describe the situation appropriately. It must be concluded that the number marking of the internal argument (horse/horses) is semantically motivated.

The assemblage-token view is clearly better able to account for the number marking of internal arguments. According to this view, an utterance with a plural subject NP is understood as conveying information pertaining to the whole assemblage, abstracting from the atomic individuals. Depending on whether the predicate involves a singular NP (cf. (5)) or a plural NP (cf. (6)), the state of affairs referred to by the whole utterance is mentally represented in terms of a relationship between the "subject"-assemblage token and one atomic token (horse) or in terms of a relationship between the "subject"-assemblage token and another assemblage token (horses). Figure 3 illustrates schematically the structure of the non-linguistic representation that, according to the assemblage-token view, is constructed for sentence (6). Because the atomic individuals of the assemblages are not represented as such, the representation does not contain any information about the exact mapping between the atomic individuals of the two assemblages. Hence, the non-linguistic representations created for (7), (8), and (9), for example, would be structurally identical to the one constructed for (6). In none of them would the relationships between the atomic individuals be spelled out (for a similar view, see Scha, 1981).

(7) At the entrance, the orphans had to show the tickets.

(8) The orphans were eating cookies.

(9) The orphans stroked the horses.

It could be argued that a coarse representation as postulated by the assemblage-token view and a fine-grained representation as postulated by the atomic-tokens view, are not mutually exclusive, but constitute representations constructed at different levels of processing during discourse comprehension. In many linguistic and psycholinguistic approaches, sentences such as (6) - (9) would be considered semantically ambiguous, because they leave unspecified the structure of the states of affairs at the level of the atomic individuals (e.g., Frazier, Pacht, & Rayner, 1999). From this perspective, it is reasonable to assume that readers or listeners attempt to resolve the ambiguity, for instance, by drawing on context information and world knowledge to eventually achieve a fine-grained representation. Accordingly, the coarse representation postulated by the assemblage-token view would be considered merely the first step on the way to a full-fledged non-linguistic representation, in which the members of the assemblages are represented by distinct tokens and the mapping between the two sets is specified.

This view presupposes that a detailed representation at the fine granularity level, in which the attributes of and interrelations among individual things are specified, is generally more appropriate than a coarse-grained representation. We doubt that this view is correct. Surely, it is often possible to spell out the exact mapping between the members of two assemblages by drawing on world knowledge and context information. Yet, it is questionable whether this is routinely done and whether constructing a fine-grained representation constitutes a deeper understanding than representing states of affairs in terms of assemblages and their interrelations. We think that for utterances with the type of plural expressions considered up to now, the intended message usually concerns assemblages. Plural expressions are used to give a rough idea of the state of affairs while ignoring irrelevant details. In a way, the matter is analogous to the issue of lexical specificity. A basic-level term (e.g., dog) is often chosen, not because specifying information is missing, but because more specific information (e.g., poodle) is irrelevant to the issue at hand and would blur the message.

Note, the claim that sentences with plural NPs (of the type considered up to now) are meant as and understood as informing about assemblages does not imply that it is generally impossible to refer to the individual members of a set. On the contrary, the claim is bolstered by the fact that there are specific linguistic devices serving exactly this purpose. An example for such a device is a singular NP involving a universal quantifier (e.g., each, every)². Sentences (3) and (10) illustrate the contrast.

(3) The orphans laughed.

(10) Every orphan laughed.

Let us assume that in both cases the discourse context is such that the listener knows that the utterance concerns the ten orphans Dr. Larch took to the circus. If it were true that proper understanding generally involves constructing a fine-grained non-linguistic representation containing distinct tokens for the individuals, the representation created for (3) would be structurally identical to the representation of (10). Yet, the utterances do not seem to be equivalent. If only 9 of the 10 orphans laughed, (3) can be accepted but not (10) (see also Dowty, 1987; Kamp & Reyle, 1993, p. 411). When uttering (3), the speaker is obviously not committed to the predicate holding for each individual. This supports the view that when a plural NP is used, the intention is not to inform about the actions or properties of the distinct individuals, but rather to inform about what holds *cum grano salis* for the assemblage. Correspondingly, a listener who constructs a fine-grained representation for (3) can be considered to have misunderstood the utterance.

Partitioning Plural NPs

The preceding considerations suggest that two cases have to be distinguished. First, the speaker considers the differences among a number of individuals to be irrelevant, represents them by an assemblage token, and denotes them by a plural NP. Second, the speaker keeps the individuals mentally apart, denotes them by using a singular NP involving a universal quantifier

(e.g., each orphan), and expresses the predicate that applies to each individual. There is, however, a third case that needs to be considered. The attribute at issue may hold for only a subset of the individuals (e.g., only 9 of the 10 orphans laughed), and the speaker may find this difference among the individuals too important to not mention it. In situations such as these, the typical means for denoting the subset of individuals that do possess the attribute are partitive plural NPs (e.g., most of the orphans) or corresponding non-partitive NPs involving a quantifier other than the definite determiner the (e.g., most orphans, few children). If it is true that grouping is precluded when the differences among the individuals are considered to be important, we expect that the VP of such a sentence expresses what holds for each individual rather than what holds for the whole subset. This indeed seems to be the case. In contrast to (11), the corresponding utterance with most of sounds odd with the plural object NP (cf. (12)). The appropriate construction would be (13), containing a singular object NP.

(11) This morning the orphans carried umbrellas on their walk to school.

(12) ?This morning most of the orphans carried umbrellas on their walk to school.

(13) This morning most of the orphans carried an umbrella on their walk to school.

The difference also becomes evident when comparing (14) and (15).

(14) The orphans gave Dr. Larch a birthday present.

(15) Most of the orphans gave Dr. Larch a birthday present.

In (14) the VP is understood as specifying what the whole assemblage did (i.e., The orphans together brought one birthday present). Yet, in (15) the very same VP is understood as describing an action performed by each individual member (i.e., Dr. Larch received several presents). Finally, the presumption that partitive plural NPs refer to what is mentally represented by distinct tokens is bolstered by analyses of the partitive determiner both, which denotes an improper subset of the background set named by the embedded NP (Reed, 1991). As illustrated in (16), both indeed seems to promote an atomic-tokens reading.

(16) For the birthday party both of the children brought a bouquet of flowers.

The same holds for simple plural NPs involving both (e.g., both children). For other simple plural NPs involving a quantifier other than the definite determiner the (e.g., most children), matters are more complicated. This type of NP seems to be associated with an atomic-tokens representation in certain contexts only. More specifically, the atomic-tokens interpretation seems to be preferred in case the context assures that the requirements of partitive NPs are met, namely that the background set is non-empty and recoverable (cf. Barwise & Cooper, 1981). For example, in (17) most orphans denotes a subset of a contextually fixed background set of orphans (the orphans living in the orphanage that Dr. Larch is responsible for). Hence, the predicate of the sentence is interpreted as applying to the individual orphans, as was the case for (15).

(17) Most orphans gave Dr. Larch a birthday present.

The main conclusion from this section is that the assemblage-token view is not valid for all types of plural NPs. In cases in which a particular subset of a given background set is denoted, the individual members are represented by distinct tokens. Specifying a subset of a background set and comparing the cardinality of the two sets for the purpose of selecting the appropriate quantifier presupposes differentiating among the members of the whole set. This may be the reason why this type of plural NP is associated with an atomic-tokens representation. In what follows we will call the plural expressions used for denoting subsets (whether proper or improper) "partitioning" plural expressions, as opposed to "non-partitioning" plural NPs which we considered in the beginning and which we assume are used for denoting assemblages.

Resolving Plural Anaphors

The first part of the paper has shown that there are two ways in which a number of individuals can be mentally represented, namely either as a single whole or as a set of distinct individuals, and that these two granularity levels of representation are associated with different

forms of NPs. It is reasonable to assume that the level of representation may be changed during discourse, depending on whether the differences among the individuals gain or lose relevance to the issue at hand. Accordingly, it is to be expected that anaphoric expressions³ do not always match their linguistic antecedents with respect to the granularity level of representation they specify. Indeed, anaphors that imply a shift in granularity are not awkward, as (18) and (19) illustrate. In (18), two individuals are introduced by a non-partitioning plural NP (two orphans), and subsequently referred to by a partitioning plural NP (both men). According to our preceding analyses, this implies a shift from an assemblage-token representation to an atomic-tokens representation.

(18) One day Dr. Larch was visited by two orphans who had left the orphanage when turning 18 years old. Dr. Larch was happy to learn that both men had found a good job.

Example (19) illustrates the opposite granularity shift. The two kittens, being introduced separately one after the other, can be assumed to be represented by distinct tokens in the beginning. In the final sentence, however, the kittens are anaphorically referred to by a non-partitioning NP (the kittens), implying a shift to an assemblage-token representation.

(19) Wally went into the backyard.

He saw a black kitten playing in the grass.

Then he noticed a white kitten watching a bird.

He wondered where the kittens came from.

Thus, anaphoric expressions do not only carry information relevant to identifying the entities the speaker refers to, but also specify the level of granularity at which they are to be represented.

Plural anaphors such as in (19) are usually referred to as anaphors with a "split antecedent." This type of anaphor has received comparatively much attention in psycholinguistic research. Murphy (1984) provided evidence that resolving a plural anaphor with split antecedent is especially time-consuming. In his study, readers needed more time to

read a sentence containing a (non-partitioning) plural anaphor such as the kittens, when the referent entities had been introduced in separate sentences (cf. (19)), than when they had been introduced by a plural NP (e.g., He saw two kittens playing in the grass. He wondered where the kittens came from). Murphy (1984) attributes the extra time needed for resolving an anaphor with split antecedent to the larger number of tokens that have to be accessed (one token for each of the two kittens for a plural anaphor with split antecedent vs. one token standing for the two kittens together for a plural anaphor with a plural NP as antecedent). According to our proposal, however, this access-based explanation is not the only possible interpretation of the finding. An anaphor as in (19) implies a shift from the atomic-tokens representation to an assemblage-token representation. Thus, to resolve the anaphor the reader must mentally group the individuals. This grouping process may partially or wholly account for the extra time needed to resolve an anaphor of this type.

It is conceivable that grouping is not always equally difficult. With nominal anaphors (e.g., the children), which were used in Murphy's (1984) study, grouping may usually be relatively easy, since it can be based on well-established categories. However, with pronominal anaphors (e.g., they), grouping may sometimes be not trivial at all. A plural pronoun may pick out heterogeneous entities, rendering grouping similar to the formation of an "ad-hoc category" (Barsalou, 1983). The following text may serve as an example.

(20) Harry polished his car lovingly.

They had had many adventures together. (Source: Kari Fraurud)

To be sure, pronominal plural anaphors with split antecedents cannot pick out any collections of entities. Linguistic analyses of the "licensing" conditions for they suggest that there has to be some correspondence among the entities denoted in the split antecedent, be it membership in the same major ontological category (e.g., humans, non-human animals, ...) and involvement in the same situation (Eschenbach et al., 1989; Fraurud, 1991; Herweg, 1988;

Rehkämper, 1991), or a common role in the current interpretive structure (Sanford & Lockhart, 1990; Sanford & Moxey, 1995). It is reasonable to assume that licensing conditions for they have something to do with how natural it is to mentally group the respective individuals. Accordingly, variables such as ontological congruency, situational relatedness, and common roles can be expected to have an impact on the processing of non-partitioning anaphoric expressions with split antecedents. This hypothesis is bolstered by experimental results. For example, according to the criteria mentioned, it should be relatively easy to group individuals introduced in an and-coordinate (e.g., Homer, Melony, and Candy went to the cinema.) or in a construction such as as well as. The reason is that in this case the individuals were already assigned "a common association base" or "a common role" through the predicate of the introducing sentence (Sanford & Moxey, 1995; see also Lang, 1984). The results of a study by Hielscher and Müsseler (1990) are consistent with this prediction. In this study, German readers were presented with sentences, each introducing two characters separately, and were asked to write continuation sentences. When the two characters were introduced in a construction like and, as well as, or neither ... nor, the participants showed a clear preference for using the plural pronoun sie (they), referring to the two characters, over a singular pronoun, referring to one of the characters. In contrast, when the two characters were introduced in constructions like without and instead of, there was a preference for a singular pronoun over the plural pronoun. Analogous results were obtained by the authors with an experimental paradigm that tapped the readers' preparedness to encounter a singular or plural pronoun in a continuation sentence (see also Albrecht & Clifton, 1998; Carreiras, 1997; Garrod & Sanford, 1982).

There are also studies in which the commonality among the individuals referred to was varied in particular respects. Carreiras (1997) investigated the impact of spatial relations. Spanish native speakers were presented with texts introducing two characters in separate clauses, and were asked to continue the text by constructing a sentence beginning with a

particular pronoun. When asked to use the plural pronoun ellos (they), response latencies were longer if the two characters had been described as being spatially far from each other than when they had been described as being close to each other. Similarly, Wilson and Stevenson (1997), analyzing reading times, found that participants read a clause containing an anaphorically used they faster when the characters referred to had been described as being together rather than apart (but see Müsseler, Hielscher, & Rickheit, 1995). A study by Sanford and Lockhart (1990) was concerned with the characters' status in the text. The authors employed a continuation task similar to the one used by Hielscher and Müsseler (1990). The status of a character was signaled by the way in which he or she was introduced in the text, either through a proper name (e.g., Harry) or through a definite description (e.g., the plumber). The two characters were either introduced in the same way or in a different way. Participants more often used a plural pronoun, referring to the two characters (rather than a singular pronoun, referring to one of the characters), when the two characters had the same rather than a different status in the text.

Taken together, the experimental findings on anaphors with split antecedents suggest that the use of they (sie, ellos) is favored when there are commonalities among the entities referred to, and that such commonalities also facilitate the reference resolution of these anaphors in language comprehension. We suppose that commonalities among entities are primarily relevant to the grouping of entities. With this interpretation, the findings bolster the view that non-partitioning expressions like they are used for and understood as referring to assemblages.

It could be argued that the findings are also consistent with a pure access explanation such as Murphy's (1984). Accessing a number of individuals may be easier when the individuals are situationally and functionally related and share a similar status in the text than when they have little in common. Thus, the effect of commonality does not necessarily reflect grouping processes, and therefore does not provide unambiguous support for the notion of assemblage-token representations.

One way to empirically decide between the access hypothesis and the grouping hypothesis, is to examine whether the commonalities among individuals also affect the resolution of partitioning anaphors with split antecedents. As was pointed out before, a partitioning expression signals to the listener that the intended level of granularity is the level of atomic individuals. Resolving such an anaphor therefore does not involve grouping processes under normal circumstances. Thus, if the commonalities among individuals are only relevant to grouping, no commonality effect should be found for the resolution of a partitioning plural anaphor with split antecedent. In contrast, if commonalities among entities facilitate access, a commonality effect should be obtained with partitioning plural anaphors with split antecedents, because accessing the individuals is obviously still necessary in these cases.

Unfortunately, experimental studies on the resolution of partitioning plural anaphors are still lacking. In the present paper we will report experiments on the resolution of both, used pronominally. According to our proposal, the pronoun both, being a partitioning expression, is understood as denoting individuals that are to be represented by distinct tokens. This view is supported by observations as to the interpretation of VPs involving internal arguments. For example, the plural object NP in (21a) appears strange, suggesting that the predicate is understood as applying to each individual. In contrast, (21b) shows that with the non-partitioning they, the same VP is interpreted differently, namely as expressing a predicate that applies to the assemblage as a whole.

(21) Homer arrived at the circus at five to eight. Cindy was already there.

- a. At the entrance, both bought two tickets for the show.
- b. At the entrance, they bought two tickets for the show.

Taken together, our proposal implies that both, in contrast to they, is understood as referring to (two) individuals, each represented by a distinct token. Accordingly, the resolution of both never requires a grouping of the individuals referred to. Commonalities among the

individuals should therefore have no impact on the resolution of both, unless they facilitate access.

Before turning to the experiments, one issue needs to be clarified. The previous considerations on access-based and grouping-based commonality effects concerned language comprehension. These considerations and the predictions derived thereof cannot directly be transferred to language production. In particular, there are reasons to suppose that in language production commonality effects may occur with both non-partitioning and partitioning expressions. Our view of the semantics of plural expressions implies that a speaker uses a non-partitioning expression (e.g., they), when considering the differences among the entities to be unimportant to the issue at hand, but uses a partitioning expression (e.g., both, all of them, each of them), when keeping the entities mentally distinct. It is likely that, by and large, the distinctness of entities is more often considered to be unimportant when the commonality among the individuals is high than when it is low, and that entities are more often kept mentally distinct, when they have little in common than when they have much in common. Thus, partitioning expressions (e.g., both) can be expected to be used relatively often in contexts that are unfavorable to non-partitioning expressions (e.g. they) (cf. Sanford & Lockhart, 1990). However, it would be wrong to consider partitioning expressions as some kind of emergency measure when entities are difficult to group. There is no doubt that partitioning expressions can be used in cases in which grouping is easy according to the criteria revealed by the above mentioned "licensing" analyses (cf. (22)).

(22) Homer and Cindy arrived at the circus at five to eight.

- a. At the entrance, both bought a bag of popcorn.
- b. At the entrance, they bought two bags of popcorn.

Moreover, a partitioning expression is sometimes even more appropriate than a non-partitioning expression in such cases. For example, both is clearly more appropriate than they in the last sentence of (23).

(23) Harry was looking for a used car and went to the best dealer in town. The first car he inspected was a two-year old Mercedes with a sliding roof. The second car was a three-year old Alfa Romeo with white leather seats. However, both were far too expensive.

These observations suggest that the choice of the referential expression does not depend on the degree of commonality or ease of grouping per se. They rather support the view that a partitioning expression (e.g., both) is used when the individuality of the entities is considered to be relevant in the given situation. This is the reason why in (23), in which two entities are juxtaposed, the last sentence needs a partitioning anaphoric expression.

In summation, we propose that non-partitioning and partitioning plural expressions serve different purposes. A non-partitioning expression (e.g., they) is used when the speaker is not interested in the differences among the entities and therefore represents the entities as an assemblage. A partitioning expression (e.g., both) is used when the speaker considers it important to differentiate between the entities and therefore represents them by distinct tokens. With respect to the listener, we assume that the different anaphoric expressions are signals as to how the entities are to be construed. In encountering a non-partitioning anaphoric expression (e.g., they) the listener has to make sure that the respective assemblage token is at hand. In case of a split antecedent, the listener must therefore mentally group the respective individuals. This process is probably easier the more commonalities there are among the individuals (grouping account). It has to be considered, however, that a high number of commonalities may also facilitate accessing the relevant tokens (access hypothesis). Thus, observing a commonality effect with sentences of this kind does not provide unambiguous support for our proposal. However, the two accounts of commonality effects can be distinguished when taking

partitioning anaphoric expressions into consideration. Our proposal implies that when encountering a partitioning anaphoric expression (e.g., both), the listener has to keep the individuals distinct. Thus, resolving a partitioning anaphor never involves grouping, and commonalities among the entities should be irrelevant. In contrast, the access account implies that commonality effects occur even with this type of anaphoric expressions, since accessing the relevant tokens is of course still necessary.

To test these ideas, two experiments were conducted. Both experiments had German native speakers as participants, and aimed at providing information about the representation of the referents of the non-partitioning anaphorically used pronoun sie (they) in contrast to the partitioning pronoun beide (both). The purpose of Experiment 1 was to examine whether our intuitions as to the difference in reference of sie (they) and beide (both) can be substantiated by the responses of a larger group of native German speakers. The participants were presented with sentence-judgment tasks similar to the VP tests employed before. In Experiment 2 reading times of sentences containing either sie (they) or beide (both) as anaphoric expression were measured. Participants were presented with short narratives in which two characters were introduced separately. Later on in the text the two characters were referred to by a pronoun. In order to obtain information as to whether reference resolution of sie (they), but not of beide (both) involves grouping the entities, ease of grouping was varied by manipulating the degree of commonality among the characters. For sake of readability the English terms they and both are used in the description of the experiments except in the materials section, where the German originals are given.

Experiment 1

The claim that non-partitioning and partitioning plural expressions differ with respect to their reference was primarily based on the impression that there is a systematic relationship between the type of subject NP (partitioning / non-partitioning) and the way the VP of such a

sentence is interpreted. Experiment 1 was to test this conjecture by assessing readers' interpretations of sentences that concerned two characters, either referred to by the non-partitioning expressions they or by the partitioning expression both. Two predictions were tested.

The first prediction concerned the interpretation of a VP which in principle is unconstrained as to whether it describes an action of an assemblage or an action of a single individual (e.g., brought a gift). Our proposal implies that a sentence with such a VP is understood as describing the action of an assemblage (i.e., only one action), if the grammatical subject of the sentence is they, but if the grammatical subject is both, it is understood as describing two "parallel" actions, performed by the two characters separately. In the experiment, the participants were presented with sentences containing VPs of this type ("unconstrained"), and for each sentence were asked a particular question that was designed to disclose the participant's interpretation of the sentence. The question was such that depending on the interpretation that was chosen, a different answer would be given. For example, when presented with the stimulus sentence (24 a) or (24 b), the participant was asked to answer the question (25).

(24) a. They brought a gift.

b. Both brought a gift.

(25) How many gifts were brought?

When interpreting the stimulus sentence as describing a joint action of the two previously introduced characters, the answer is most likely "one." However, when interpreting the sentence as describing two "parallel" actions of the two characters, the answer is most likely "two." Thus, as to this particular item, our hypothesis predicts that participants answer "one" when being presented with the they-version of (24), but "two" when being presented with the both-version of (24). For the other items, analogous diagnostic questions were employed. In order to obtain

information as to whether sentences that were interpreted contrary to our prediction were found awkward, participants were also asked to rate the acceptability of the sentences.

The second prediction concerned predicates that can only be applied to assemblages, i.e., collective predicates (e.g., 'met at the cinema') and predicates which according to world knowledge usually do not apply to single individuals, (e.g., 'received the diplomas from the principal's hand'). Our proposal implies that such a predicate can only be combined with a non-partitioning term (e.g., they) as grammatical subject. Sentences containing this type of VP together with both as subject should therefore be judged as unacceptable. In order to test this prediction, participants were presented with another set of sentences containing they or both as grammatical subject combined with this type of VP ("constrained"). With these sentences, participants were only asked for acceptability ratings. If our hypothesis is correct, the both-version of these sentences should be rated less acceptable than the they-version.

Method

Participants. Sixty students and faculty members of various disciplines took part in the experiment. They were randomly assigned to two groups of 30 participants each. All participants were native German speakers.

Materials. Materials comprised 18 written items (see Appendix for a complete list)⁴. Items were simple German sentences. Each item was available in two versions, with the grammatical subject being either sie (they) or beide (both). As the participants were told in the beginning, all sentences were to be interpreted as referring to the same two characters, Fritz and Maria. The VP of the sentences involved a verb with a direct object, a reflexive verb, or an intransitive verb.

For the first set of 9 sentences (unconstrained), the VP was such that when considered in isolation, it could be read as expressing either the action of an assemblage or the action of a single individual (e.g., brachten ein Geschenk mit (brought a gift)). Depending on the reading chosen, however, the described state of affairs differed in one specific aspect (either one or two

gifts are brought). This aspect was addressed by a question written below the stimulus sentence (e.g., Wie viele Geschenke wurden mitgebracht? (How many gifts were brought?)), followed by an empty space for the answer. In addition, participants were asked to rate the acceptability of the sentence, using a rating scale ranging from 1 (perfectly acceptable) to 5 (not acceptable at all), which corresponds to the grading system in German schools.

For the second set of 9 sentences (constrained), the VP expressed a predicate which, for logical reasons or according to world knowledge, is applicable to assemblages only (e.g., trafen sich am Kino (met at the cinema)). With these items, participants were asked for acceptability ratings only.

All 18 items were put together into one questionnaire, with the unconstrained set appearing first and the constrained set appearing second. There were two versions of the questionnaire, A and B, which were complementary to each other with respect to the selections of items presented in the sie (they) version and the beide (both) version. Assignment was such that each questionnaire contained 9 sie (they) items and 9 beide (both) items (Questionnaire A: 5 unconstrained and 4 constrained sie (they) items, 4 constrained and 5 unconstrained beide (both) items. Questionnaire B complementary).

A brief paragraph at the beginning of the questionnaire informed that all of the following items concerned the two characters Fritz and Maria. The paragraph also explained the acceptability rating task.

Procedure. Questionnaires were sent via e-mail to the participants, with one group of participants receiving the A version and the other group receiving the B version.

Scoring. Responses to the questions of the unconstrained set were scored by two independent judges naive to the hypothesis of the experiment. The judges were instructed to determine whether a given response indicated that the participant interpreted the sentence as describing one action in which both characters were involved (assemblage reading) or two

parallel actions (atomic reading). Judges were told that if a response indicated that both readings were noticed, the response was to be scored with respect to the prominence of one or the other reading. More specifically, five interpretation categories were to be distinguished: A response was to be scored 1 if it indicated that only the assemblage reading had been taken into consideration (e.g., participant responds with "one" to the question "how many gifts were brought" after reading (24a) or (24b)), it was to be scored 2 if the assemblage reading was mentioned first but the atomic reading was also mentioned (e.g., the participant responds with "one, maybe two"), the response was to be classified as 3 if the response indicated that both readings had been recognized but did not give a hint as to which reading had been thought preferable (e.g., the participant responds with "one or two, both equally plausible"), the response was to be scored as 4 if the atomic reading was mentioned first but the assemblage reading was also mentioned (e.g., the participant responds with "two, maybe one"), and finally the response was to be scored as 5 if only the atomic reading was mentioned (e.g., the participant responds with "two"). Missing responses as well as responses off the issue were to be classified as blanks. Blanks were excluded from further analyses (2.96%). There was 97.8% agreement among the judges. Discrepancies were subsequently resolved in a joint discussion. For each participant and each item, respectively, the mean score in each pronoun condition was determined. These means will be referred to as "interpretation scores."

Results

In what follows F_1 and t_1 refer to analyses with error terms based on participant variability, and F_2 and t_2 refer to analyses with error terms based on item variability.

Unconstrained set. The mean interpretation score was $M = 1.40$ with they-sentences, indicating a preference for the assemblage reading, and $M = 3.50$ with both sentences, indicating a bias towards the atomic reading. This difference proved highly significant according to the t

test for correlated samples, $t_1(59) = 17.53$, $p < .01$; $t_2(8) = 4.78$, $p < .01$. The frequencies of the various interpretation categories for the two types of sentences can be found in Table 1.

The mean acceptability ratings are also shown in Table 1. For they-sentences, the acceptability ratings were most favorable (i.e., lowest grades were assigned) if the sentence was unambiguously interpreted in the predicted way (assemblage-reading only), and became less and less favorable the more the atomic interpretation came forth (Interpretation Category 2 to 5). For both-sentences, the pattern was less clear, although the ratings were still slightly more favorable for sentences that were interpreted in the predicted way (atomic interpretation) than for sentences interpreted in another way. Since there were no participants and only two items with observations in all 10 conditions, the acceptability ratings could not be submitted to repeated measurement analyses. Instead, a participant's mean ratings in the different conditions were treated as independent observations, as were the mean ratings for a given item in the various conditions. Accordingly, ANOVAs were performed that involved two between factors, pronoun (they, both) and interpretation category (Category 1 - 5), with unequal numbers of observations in the cells. The main effect of pronoun (they / both) on the acceptability ratings was not significant, $F_1(1,222) < 1$; $F_2(1,57) < 1$, but the main effect of interpretation was, $F_1(4,222) = 7.48$, $p < .01$; $F_2(4,57) = 3.27$, $p < .05$. More important, there was an interaction of pronoun and interpretation, which however was only significant in the analysis by participants, $F_1(4,222) = 7.99$, $p < .01$; $F_2(4,57) = 1.45$, $p > .10$. Breaking down the interaction yielded the following results. Separate analyses for each interpretation category showed that they- and both-sentences differed significantly in acceptability for two of the interpretation categories. As expected, with Category 1 (assemblage only), they-sentences were rated as more favorable than both-sentences, $t_1(103) = 4.69$, $p < .01$; $t_2(16) = 2.98$, $p < .01$, while with Category 5 (atomic only), both-sentences were rated as more favorable than they-sentences, $t_1(66) = 4.03$, $p < .01$; $t_2(12) = 1.37$, $p < .10$ (one-tailed). Separate analyses of the interpretation effect for the two sentence

types confirmed that interpretation affected the acceptability ratings for they-sentences, $F_1(4,96) = 12.29$, $p < .01$; $F_2(1,25) = 2.90$, $p < .05$, with the linear trend being significant, $F_1(96) = 33.10$, $p < .01$; $F_2(1,25) = 6.21$, $p < .05$. With both-sentences, however, the effect of interpretation was not statistically significant, $F_1(4,126) = 1.57$, $p > .10$; $F_2(4,32) = 1.22$, $p > .10$.

Constrained set. For each participant and for each item, mean acceptability ratings were determined in the they- and the both-versions, respectively. These mean scores were submitted to t -tests for correlated samples. As predicted, sentences were assigned significantly better acceptability grades when the grammatical subject was they ($M = 1.3$) than when it was both ($M = 2.1$), $t_1(59) = 9.33$, $p < .01$; $t_2(8) = 4.41$, $p < .01$.

Discussion

The purpose of the experiment was to provide information as to whether they, being a non-partitioning expression, is usually interpreted as referring to an assemblage of entities, whereas both, being a partitioning expression, is usually interpreted as referring to two distinct entities. The results bolster this hypothesis. In the unconstrained sentence set, the predicate of the sentence was most often interpreted as applying to the two characters together when referred to by they, but as applying to each of the characters separately when referred to by both. Moreover, in those cases in which participants did not interpret the predicate in this way, they rated the sentence's acceptability as relatively low. The result of the constrained sentence set confirmed that combining both with a VP that expresses a predicate that applies to assemblages only, yields less acceptable sentences than combining they with such a VP. It should be noted that this difference cannot be attributed to both-sentences being less acceptable in general, because there was no overall difference in acceptability between both- and they-versions in the unconstrained sentence set. Moreover, both-sentences were rated as more acceptable than they-

sentences when the predicate was understood as applying to each of the atomic individuals separately.

Experiment 2

According to our hypothesis, a reader who encounters the anaphoric expression they, referring to individuals that were introduced separately, must mentally group the individuals and construct an assemblage token. This grouping process is assumed to vary in difficulty depending on the degree of commonality among the individuals. Accordingly, the time needed for reference resolution of they, if referring to separately introduced individuals, can be expected to depend on the degree of commonality between the individuals being referred to. For both, however, the prediction is different. Since both is a partitioning expression, our hypothesis implies that it does not call for grouping the individuals referred to. Accordingly, the time needed for reference resolution of both should not depend on the degree of commonality between the individuals being referred to.

These predictions are based on the assumption that commonality solely affects grouping. As was pointed out in the first part of the paper, however, commonality may also have an impact on accessing the individuals being referred to. A number of individuals may be easier to access if they have a lot in common than if they have little in common. If commonality indeed affects access, a commonality effect should occur not only with the resolution of they, but also with the resolution of both, because accessing the individuals referred to is necessary with both anaphoric expressions. Depending on whether or not it is assumed that commonality affects access in addition to grouping, more specific predictions can be derived. If commonality is only relevant to access, then it should affect the resolution of they and both to a similar extent. In contrast, if commonality affects access in addition to grouping, then there should be a commonality effect for both pronouns, with the effect being stronger for they than for both. It should be noted that such a commonality by pronoun interaction would not be inconsistent with

our proposal. It would suggest that although sharing access-based effects, the reference resolution of they involves grouping the individuals referred to whereas the resolution of both does not. However, for ease of presentation, we will consider only the pure grouping hypothesis when deriving predictions from our proposal in what follows.

In the experiment, participants were presented with short narratives, sentence by sentence (self paced). In each narrative, two main characters were introduced one after the other in separate sentences. The degree of commonality between the characters was varied. More specifically, there were four versions of each text, differing with respect to the description of the second character's action goal and spatial distance from the first character (see sample text in Table 2). In one version the characters had little in common; the second character pursued a different goal and was spatially far from the other character ("different-far"). In a second version, the second character cooperated with the first character, pursuing the same goal, though being spatially far away ("same-far"). In a third version, the second character pursued a different goal but was spatially close to the other character ("different-close"). In a fourth version the two characters cooperated and were close to each other ("same-close"). The text segment describing the second character was followed by the target sentence, which contained a plural pronoun unambiguously referring to the two characters. The plural pronoun was either sie (they) or beide (both). Degree of commonality and type of pronoun were crossed, amounting to altogether eight versions of each text (i.e., 4(commonality) x 2(pronoun)).

If it is true that readers interpret they but not both as referring to an assemblage, and that increasing the degree of commonality is relevant to assemblage-token construction only, reading times for the target sentences containing sie (they) should vary depending on the degree of commonality, whereas reading times for the target sentences containing beide (both) should not. In contrast, if degree of commonality is relevant to accessing the tokens standing for the two characters, a commonality effect should also obtain with target sentences containing both.

It has to be taken into consideration, however, that assemblage-token construction may not only be triggered by they (or other non-partitioning referential expressions), but also by the predicate of a sentence, in case the predicate is of the "constrained" type (i.e., applicable to assemblages only). To minimize problems as to the interpretation of the expected commonality effect, all predicates were instances of so-called distributive predicates (e.g., 'to smile', 'to look forward to', 'to know').

Our proposal also allows for certain predictions concerning sentences following the target sentences. The notion of assemblage tokens implies that once an assemblage token has been constructed it is available for the resolution of subsequent anaphoric expressions. That is, having processed a sentence containing a non-partitioning plural expression (e.g., they) may have consequences for processing the subsequent sentence. If the subsequent sentence again contains a non-partitioning plural expression (e.g., they) referring to the same individuals, the appropriate assemblage token is already available. It can simply be picked out and linked to the new predicate. No grouping is necessary. Thus, the reading time for the second sentence should not be affected by the degree of commonality. In contrast, if the first sentence contained a partitioning plural expression (e.g., both) and the subsequent sentence contains a non-partitioning plural expression (e.g., they), grouping has to be made up for when processing the second sentence. Thus, in this case, the reading time for the second sentence should be affected by the degree of commonality. To test these predictions, every target sentence of this experiment was followed by a sentence that involved the anaphoric pronoun they referring to the same characters as the target sentence. This sentence ("post-target sentence") was the same in all eight text versions. Thus, the post-target sentence had they as grammatical subject not only for those participants who had received they in the target-sentence ("they-they condition") but also for those participants who had received both in the target sentence ("both-they condition"). In the they-they condition, resolving the they in the post-target sentence should not require grouping

because an assemblage token standing for the two characters should already be available from resolving the anaphor they in the target sentence. In contrast, in the both-they condition, resolving they in the post-target sentence should require a grouping of the characters, because grouping did not take place when processing the target sentence.

Taken together, we expect the degree of commonality to affect the reading times for target sentences in the they-they condition, and for post-target sentences in the both-they condition. We do not expect the degree of commonality to affect reading times for the target sentences in the both-they condition, or the post-target sentences in the they-they condition.

Method

Participants. Sixty-four students of the Technical University of Berlin participated in the experiment for course credit. They all were native speakers of German.

Materials. Materials comprised 32 German narrative texts. There were 16 experimental texts, constructed according to the following scheme (see Table 2). First, the setting of the story was outlined, and the action and location of the first character, who was always female, were described. In the next section, the action and location of the second character, who was male, were specified. In this section the degree of commonality between the two characters was varied: The second character was described as either pursuing the same goal as the first character (same) or as pursuing his own goals which were independent of the goals of the first character (different). Also, the second character was located either in range of sight of the first character (close) or in a different building (far). Although the variation of the degree of commonality was based on two variables (goal and distance), we did not consider it a two by two variation since the text versions did not realize a strict orthogonal variation of the two factors. Strict orthogonal variation would have brought about implausible, incoherent texts. Yet, care was taken to keep the four versions as similar as possible with respect to the concepts mentioned and the relatedness of the two characters in the propositional structure of the text.

Prior to the target sentence, the two characters were never referred to in the same sentence. There were two versions of the target sentence, which were crossed with the four commonality versions. The two versions of the target sentence differed only in the plural pronoun, which was either sie (they) or beide (both). The pronoun was in grammatical subject position and unambiguously referred to the two characters in all conditions. The VP of the sentence described an action or activity that is typically performed by a single individual (e.g., 'to smile', 'to look forward to', 'to know'). The next sentence (post-target sentence) contained the plural pronoun sie (they) as subject, referring to the same two characters that were referred to in the target sentence. This sentence was the same in all eight text versions.

The 16 filler texts were of comparable length and concerned similar topics. They described actions of one, two, or three characters.

To ensure that the participants were reading for comprehension, each text was followed by a declarative sentence concerning information conveyed by the text, which had to be verified by the participants. For half of the texts this sentence was true and for the other half it was clearly false given the information provided by the respective story.

Design and Procedure. Thirty-two of the participants were presented with the text versions containing sie (they) in the target sentence (they-they condition), and the other 32 participants were presented with the text versions containing beide (both) in the target sentence (both-they condition). Each of the two groups of participants was further subdivided into four groups of eight participants each, and the 16 experimental texts subdivided into four sets. Commonality versions were assigned to participants and texts according to a 4(group) x 4(set) x 4(commonality) Latin square, within the they-they condition as well as within the both-they condition.

Each participant was presented with the 16 experimental and the 16 filler texts in a different random order. Texts were presented sentence by sentence on the computer screen.

Participants used the space bar to advance through the text. When pressing the space bar after having read the final sentence of a text, the corresponding verification sentence appeared on the screen marked by the heading "Behauptung" (statement). Participants used the d or l key on the keyboard for the no or yes response, respectively. Participants were instructed to read carefully but at a normal pace. Each session began with two examples to make sure that participants understood the procedure. The entire session lasted approximately 40 minutes.

Results

In determining outliers, it was taken into account that there were considerable differences among the 32 critical sentences (16 items x 2 sentences) with respect to length and complexity, and consequently also with respect to the mean reading time. Reading times that deviated more than two standard deviations from the mean reading time for the respective item in the respective condition were excluded from further analyses. This eliminated 5.3 % of the reading times. In order to take the sentence-length variability out of the error terms, the linear regression of sentence reading time on the number of characters of a sentence was computed for each participant separately, and the residuals were submitted to further analyses (cf. Ferreira & Clifton, 1986). Figure 4 displays the mean residuals in the four commonality versions of the target and post-target sentences for the they-they condition and the they-both condition.

The residuals were submitted to two mixed ANOVAs with participants and items as random variables, respectively. Participants and items were assigned to groups and sets, respectively, according to the corresponding dimension of the above-mentioned Latin square. The analysis by participants involved a 2(pronoun: they-they vs. both-they) x 4(group) x 2(sentence: target vs. post-target) x 4(commonality) design with pronoun and group as between variables, and sentence and commonality as within variables. The analysis by items involved a 4(set) x 2(pronoun: they-they vs. both-they) x 2(sentence: target vs. post-target) x 4(commonality) design with set being the only between variable. However, since the analysis

did not yield any significant effects of set (all $F_s < 1$), this variable was dropped from the analysis (Pollatsek & Well, 1995).

The main effect of sentence was significant in the analysis by participants but not in the analysis by items, $\underline{F}_1(1,56) = 72.19$, $p < .01$; $\underline{F}_2(1,15) = 2.57$, $p = .13$. The same held for the degree of commonality, $\underline{F}_1(3,168) = 3.84$, $p < .05$; $\underline{F}_2(3,45) = 2.45$, $p = .08$. The main effect of pronoun was not significant, nor were the two-way interactions, all $\underline{F}_s < 1$. Most important, however, the three-way interaction of pronoun, sentence, and commonality proved significant, $\underline{F}_1(3,168) = 3.30$, $p < .05$; $\underline{F}_2(3,45) = 2.84$, $p < .05$. To determine the basis of this interaction, the effect of commonality was tested for each of the four pronoun by sentence conditions separately. As predicted, the effect of commonality proved significant with target sentences in the they-they condition, $\underline{F}_1(3,84) = 4.22$, $p < .01$; $\underline{F}_2(3,45) = 3.03$, $p < .05$, as well as with post-target sentences in the both-they condition, $\underline{F}_1(3,84) = 2.63$, $p = .055$; $\underline{F}_2(3,45) = 2.95$, $p < .05$. The effect of commonality was not significant with target sentences in the both-they condition, $\underline{F}_1, \underline{F}_2 < 1$, nor with post-target sentences in the they-they condition, $\underline{F}_1(3,84) = 1.12$, $p > .30$, $\underline{F}_2 < 1$. Analyzing the significant simple effects of commonality by means of the Duncan test yielded the following results. With target sentences in the they-they condition, reading times were significantly longer in the same-far condition than in both, the same-close condition, p_1 and $p_2 < .05$, and the different-close condition, p_1 and $p_2 < .05$. The difference between same-close and different-far approached significance, p_1 and $p_2 < .10$. With post-target sentences in the both-they condition, reading times were significantly longer in the same-far condition than in each of the other conditions, all p_1 and $p_2 < .05$.

Discussion

The results of the experiment support the view that resolving the non-partitioning pronoun they requires an assemblage token, whereas resolving the partitioning pronoun both does not. Degree of commonality, which is assumed to be relevant to assemblage-token construction, had

a significant effect on the reading times for the target sentences containing they, but did not affect the reading times for the target sentences containing both. The findings pertaining to the post-target sentences also supported this view. First, in the they-they condition, reading times for post-target sentences were unaffected by the degree of commonality. This is consistent with the hypothesis that the readers already grouped the characters into an assemblage when processing they in the target sentence, and thus did not need to group them again when processing the post-target sentence. Second, the degree of commonality did affect the reading times for post-target sentences in the both-they condition. This result supports the hypothesis that the readers did not group the characters into an assemblage when processing both in the preceding target sentence, and thus still needed to do so when processing the post-target sentence. In short, the pattern of the results for the effect of commonality can be accounted for by our proposal regarding the representation of the referents of they and both, together with the assumption that the degree of commonality among the individuals referred to has an impact on grouping them into an assemblage.

The more specific results for the four commonality conditions, however, suggest that matters might be more complicated. In those conditions in which commonality had a significant effect, reading times were not longest with the different-far condition, as one might have expected. Different-far produced shorter mean reading times than same-far, and with the post-target sentence of the both-they condition, different-far even produced the shortest mean reading times of all four commonality conditions. Considering that this latter condition was the most disadvantageous one for grouping, it may be speculated that readers dispense with creating an assemblage token for they if most of the information conveyed by the context speaks against grouping the individuals. When the post-target sentence is presented after a different-far text version and a target sentence with both, not only the spatial distance and the divergent goals but also the partitioning expression (both) in the immediately preceding target sentences stresses the

distinctness of the two characters. In such an extreme condition, the linguistic command conveyed by they might simply not be followed by the readers.

The results of the experiment do not provide positive evidence for the hypothesis that the degree of commonality is relevant to accessing the individuals referred to. Had the access hypothesis been correct, a commonality effect should have emerged not only with the target sentences containing they, but also with the target sentences containing both, since accessing the individuals referred to is necessary in both cases. Yet, there were no significant differences among the four commonality conditions for the reading times of target sentences containing both.

This negative result appears surprising when considering that research on the resolution of singular anaphors has provided evidence that anaphor resolution times are influenced by the commonality between the protagonist of a story and the entity referred to. This has been shown especially clearly for commonalities with respect to spatial location (Glenberg, Meyer, & Lindem, 1987, Müsseler et al., 1995; Rinck & Bower, 1995). Since singular anaphors do not require the construction of an assemblage token, these commonality effects cannot be attributed to grouping processes, but only to access processes. A possible resolution of the discrepancy is related to the different roles the referents play in the texts. In text comprehension research, access-based effects of spatial location are usually explained in terms of focus shifting. It is assumed that space is represented analogously at the non-linguistic level of representation, and that a space-based "spotlight" is focused on the protagonist's location. Accessing entities at other locations takes time, since the spotlight has to be shifted accordingly (Bower, 1989; Morrow, Bower, & Greenspan, 1989, 1990; Morrow, Greenspan, & Bower, 1987). In the studies on singular anaphors, the referent of the anaphoric expression was an entity that was located either nearby or far away from the protagonist of the story, i.e., the entity was either within or outside of the scope of the spotlight. In our texts, however, the entities referred to by the plural

pronoun were the two protagonists of the story. In this case the spotlight was probably adjusted so that both protagonists were within its scope (cf. "neutral perspective" in Franklin, Tversky, & Coon, 1992). Thus, there was no need of shifting the spotlight.

On this interpretation the discrepancy of results is not related to the distinction between singular and plural anaphors but rather to the location of the target entities in relation to the spotlight. Extending this idea to other (i.e., non-spatial) dimensions of commonality, leads to the hypothesis that access-based commonality effects only arise with target entities outside of the focus of attention. Accordingly, we would expect an access-based commonality effect to occur with a plural anaphor that refers to non-essential objects outside of the current focus of attention. Conversely, we would expect that the usual spatial-distance effect established for singular anaphors does not emerge if the singular anaphor refers to one of several main characters. Future research has to show whether these predictions are correct.

Our experiment was designed to investigate the impact of commonality on anaphor resolution. No predictions were made with respect to the mean reading times for target and post-target sentences in the two pronoun conditions (they-they vs. both-they). Yet, some of the respective results may appear surprising and warrant a comment. The mean reading times for target sentences containing both were not shorter than those for target sentences containing they. At first glance, this result may appear inconsistent with the claim that they but not both requires grouping of the individuals. It should be noted, however, that according to our proposal, processing a sentence containing both does not merely require accessing the appropriate atomic tokens but also assigning the predicate to each of the two tokens separately (see also Murphy, 1984). Assigning the predicate to two instead of one token may take at least as much time as grouping the individuals would. It could be argued that this assumption still leaves unexplained why mean reading times for the target-sentences containing both were even longer than the reading times for their respective post-target sentences. However, it should be kept in mind that

different sets of sentences were employed as target and post-target sentences. The difference in reading time may therefore simply be due to uncontrolled differences in the complexity of the sentences' linguistic structure and content. For the same reason we are also reluctant to draw conclusions from the analogous target/post-target difference with the they-they condition.

General Discussion

The linguistic observations presented in the first part of the paper and the experimental results reported in the second part provide converging evidence for the hypothesis that there are (at least) two ways in which a number of individuals can be mentally construed, and that there are different linguistic expressions used for denoting the respective individuals in the two cases. The VP tests that we employed in the first part of the paper revealed that plural NPs do not form a homogenous category with respect to the nature of the mental representation of their referents. A "standard", non-partitioning plural NP seems to be used for and understood as referring to individuals mentally construed as a single whole, as an assemblage. In contrast, a partitioning plural NP, which "picks out" a subset from a given background set (e.g., most of the orphans, most orphans (of one particular orphanage)), is tied to a representational structure which contains distinct tokens for the individual members. This difference between non-partitioning and partitioning expressions is also reflected in the results of the experiments. Experiment 1 showed that the predicate of a they-sentence is primarily understood as applying to the two characters together, whereas the predicate of a both-sentence is primarily understood as applying to each of the individual characters separately. The results of Experiment 2 confirm that the degree of commonality between the characters referred to is relevant to reference resolution of they (when encountered for the first time) but not for reference resolution of both. Taken together the experimental results suggest that the pronoun they refers to what is conceived of as an assemblage, while both refers to what is conceived of as two individual

entities. This supports the idea that the distinction between assemblage-token representations and atomic-tokens representations has its linguistic counterpart in the distinction between non-partitioning and partitioning plural expressions.

The contrast between they and both

It could be argued that sie (they) and beide (both) differ not only with respect to the non-partitioning/partitioning dichotomy, but in several other respects as well (see Reis & Vater, 1980, and Schopp, 1995, for a detailed analysis of beide). The differential results obtained for sie and beide may therefore be due to some other difference, unrelated to our proposal. However, we do not consider this a convincing objection. First, in the experiments, sie as well as beide were used in only one particular function, namely as pronominal anaphoric expressions. Usages that are possible for sie, but not for beide, or vice versa, were excluded. More specifically, sie was never used deictically, and beide was never used pronominally (e.g., beide Kinder (both kids), die beiden Kinder (the two kids)), or as a floating quantifier (Die Kinder waren beide hungrig (The children were both hungry)). Furthermore, in the present experiments, sie --just as beide-- always referred to only two characters, and this was clear from either the explicit instruction (Experiment 1) or the restricted discourse domains of the narratives (Experiment 2). Hence, if at all, it could be objected that differences in the potential usages of sie and beide may have contributed to the results of the experiments. Second, and more important, we did not just obtain any differences between sie and beide, but particular differences with respect to the interpretation of verb phrases (Experiment 1) and the impact of the entities' commonality on resolving the anaphor (Experiment 2). These differences were predicted on the basis of the partitioning/non-partitioning distinction. It is difficult to see how they could be accounted for on the basis of any of the other differences between sie and beide.

Thus, it seems well justified to attribute the differential results obtained for sie and beide to the fact that sie is a non-partitioning expression whereas beide is a partitioning expression.

Similarities and differences between our proposal and formal-semantic theories of plurals

According to our proposal, atomic-tokens and assemblage-token representations constitute alternative ways of conceptualizing multiple entities. More specifically, if the differences among entities appear unimportant to the issue at hand, the entities are represented in terms of an assemblage. The corresponding linguistic means for denoting the entities is a non-partitioning plural NP. In contrast, if differentiating between the entities is considered to be important, they are mentally represented by distinct tokens and denoted by a partitioning expression. As was pointed out in the first part of the paper, there are several formal-semantic theories of plurals that postulate complex individuals in addition to atomic individuals (e.g., Kamp & Reyle, 1993; Landman, 1989; Link, 1983; 1991; Scha, 1981; Schwarzschild, 1994; Verkuyl & van der Does, 1996; Winter, 2000). In this sense, our notion of assemblage tokens is not new. However, in contrast to our proposal, most semantic theories of plural do not tie the complex/atomic-distinction to the form of the plural expression, but rather to the predicate of the sentence containing the plural expression as grammatical subject (for mixed proposals, see below). A predicate that only allows for a collective interpretation (e.g., 'to gather') is always applied to a complex individual. In contrast, a predicate that is considered to only allow for a distributive interpretation (e.g., 'to laugh'), is applied to each atomic individual separately. Finally, when a predicate allows for both interpretations (e.g., 'to carry a table') the sentence is considered to have two readings, corresponding to the two interpretations of the "mixed" predicate. To be a little more precise, the logical form postulated for a plural definite description generally contains a variable for a complex individual (e.g., X for the denotation of the men). In case the plural expression is combined with a collective predicate, the predicate is directly applied to this complex individual (e.g., gather(X)). In case of a distributively interpreted predicate, however,

the complex individual is taken apart and the predicate is applied to each of the atomic individuals that together make up the complex individual (e.g., $\forall x \in X[\text{smile}(x)]$). Thus, in any case, the logical form contains a variable for a complex individual, but depending on the predicate, the logical form also contains a variable for atomic individuals.

According to our proposal, it is solely the type of plural NP, not the type of predicate that determines whether a number of distinct tokens or an assemblage token is used for representing the denotation of the NP.⁵ Thus, even if the predicate is one that would clearly be classified as distributive (e.g., 'to laugh'), the denotation of the plural NP is represented as an assemblage if the NP is a non-partitioning plural expression (e.g., the children in the children laughed). At first glance, combinations of such predicates with assemblages may appear odd, rendering our proposal less plausible than the formal-semantic proposals, which explicitly rule out these combinations. However, it should be kept in mind that our assemblage/atomic tokens distinction concerns the grain-size of representations. When proposing that the children laughed is associated with a representation in which the property 'laughing' is assigned to an assemblage of children, we do not deny that the speaker or listener in principle knows that the children's laughing is made up from the laughing of individual children. We merely claim that at this particular moment, the speaker or listener does not adopt such an analytic view, but employs a representation in which the exact properties of the atomic individuals are not spelled out (for a similar view see Scha, 1981).

Our proposal, according to which the denotation of a (non-partitioning) plural NP is generally represented as an assemblage, has a number of advantages. With our proposal certain problems do not arise that are plaguing semantic theories that tie the individual/complex distinction to the distributive/collective distinction. First, if it is assumed that distributive predicates are generally associated with logical forms involving atomic-individuals, a sentence

like The children laughed must be considered as being equivalent to Every child laughed. However, this view is questionable (Dowty, 1987; Winter, 1998).

Second, the traditional distributive/collective dichotomy is probably too simple (cf. Verkuyl & van der Does, 1996; Winter, 2000). There are predicates (or readings of predicates) that seem to be neither distributive nor collective in the usual sense (e.g., 500 Dutch firms own 6000 American computers, Scha, 1981).

Third, it seems difficult to justify why predicates such as 'to laugh' or 'to smile' should not be combinable with complex individuals. Take, for example, the predicates 'to have an idea' and 'to call a taxi.' Having an idea is a cognitive event happening inside an individual's mind, and similarly, calling a taxi is an act that is usually performed by one individual alone (there is usually only one individual who actually makes the phone call). Hence, according to the usual criteria, both predicates are clearly distributive predicates, comparable to 'to laugh' or 'to smile'. Sentences such as The children had an idea how to solve the problem or The men called a taxi should therefore be analyzed as having only one reading, namely the distributive reading. Yet, most formal proposals would ascribe these sentences not only the distributive reading, but also the collective reading, taking into account the intuition that the sentences each may describe only one event (i.e., the children may have figured out the solution together; one taxi was called). Although this concession is reasonable from an empirical point of view, it is difficult to see how the collective reinterpretation of the distributive predicate could be justified theoretically for cases such as these, without conceding the possibility for other distributive predicates as well (e.g., 'to laugh', 'to smile').

From a psycholinguistic point of view there appears to be an additional problem, concerning the notion of mixed predicates (e.g., 'to carry a table', 'to bring a gift'). In most semantic theories, a sentence containing a (non-partitioning) plural subject NP and a mixed predicate is ascribed two readings, a collective reading and a distributive reading. Yet,

Experiment 1 showed that with sentences such as They brought a gift the one-act reading (i.e., collective reading) is clearly preferred. Similarly, an eye-movement study by Frazier, Pacht, and Rayner (1999) showed that when processing a sentence with a conjoined NP and a mixed predicate (e.g., Juli and Jason baked several cakes), the collective reading is chosen, unless evidence for the distributive reading (e.g., a floating each) is encountered. Of course, the results of these studies could be dismissed as irrelevant to semantic theories, as they do not show that native speakers strictly reject the distributive reading of these sentences. However, it may be more fruitful to take findings such as these as an occasion to reflecting upon the proper semantic analysis of such sentences.

Our findings may be of interest for the semantics of plurals in still another respect. In most formal-semantic proposals, the form of the plural expression is not considered relevant to distinguishing between cases in which the predicate is applied to a complex individual and those in which the predicate is applied to atomic individuals. However, there are exceptions (Kamp & Reyle, 1993; Winter, 1998). For example, Kamp and Reyle (1993) posit that the logical form of a non-quantifying plural expression (e.g., the children) generally contains a variable for a complex individual (which will be taken apart when combined with a distributive predicate). In contrast, the logical form of a quantifying expression (e.g., most children) is assumed to only contain a variable for atomic individuals. Thus, as we do, Kamp and Reyle propose that an expression such as most of the orphans or both orphans does not trigger the constructing of a token standing for a complex individual. The results of our study suggest that a distinction along these lines is indeed needed. The two types of plural expressions give rise to systematically different interpretations, as was shown for nominal expressions in the first part of the paper, and for pronouns in the second part. Moreover, our experimental results suggest that processing a non-partitioning expression such as they, in contrast to processing a partitioning expression such as both, is affected by variables that can be considered relevant to assemblage token

construction. These differences were found even though all predicates used in the texts were what proponents of formal-semantics proposals would class as distributive. These findings therefore support the view that the relevant variable is the nature of the plural expression rather than the nature of the predicate.

The bound-tokens view

We have limited considerations to two ways of representing referents and the predicates ascribed to them. However, depending on the text comprehension model that is chosen, other representational views appear possible as well. Let us briefly discuss a representational structure that suggests itself when a conceptual network model is chosen. In a conceptual network model, a statement about a single individual (e.g., Peter laughed loudly) is typically assumed to be represented by mapping the token node, standing for the referent, onto the concept nodes representing the predicate (e.g., Glenberg & Langston, 1992). Accordingly, for assertions concerning more than one individual, the most obvious assumption is that the token nodes, standing for the individuals referred to, are each mapped onto the node or nodes representing the predicate of the sentence. In more general terms, the assumption is that the individuals are represented by distinct atomic tokens, but the predicate is represented only once, binding the atomic tokens together (see Figure 5). We will refer to this representational view as the “bound-tokens view.”

In certain respects a bound-tokens representation resembles an assemblage-token representation. Although the atomic tokens are not grouped into an assemblage token in a bound-tokens representation, the tokens are still in some sense grouped, namely by being mapped onto the same predicate representation. Collective predicates can therefore easily be accounted for by the bound-tokens view. In fact, the proposal by Sanford and Moxey (1995), which states that the referents of they and other (in our terms non-partitioning) plural NPs are mapped into a common role slot (see also Murphy, 1984), comes very close to a bound-tokens

view. Yet, there are two important differences between the assemblage-token view and the bound-tokens view. The first difference concerns the stability of the grouping of the individuals. Once constructed, an assemblage token is permanent in principle. If an anaphoric expression in a subsequent sentence refers to the same assemblage of individuals, the atomic tokens do not have to be accessed again. Rather, the already available assemblage token can be picked up and mapped onto the predicate expressed in the sentence. In contrast, in a bound-tokens representation, the binding of the tokens is accomplished by the particular predicate given in the sentence, and therefore provides merely a temporary grouping (for a similar distinction see Grosz & Sidner, 1986). If in a subsequent sentence with a new predicate a plural expression refers to the same individuals, the atomic tokens have to be accessed again and mapped onto the new predicate. Thus, a set of tokens bound together by a predicate is not functionally equivalent to an assemblage token with regard to processing repeated plural reference. The second difference concerns the status of the grouping process itself. Grouping of the individuals is a necessary step in creating an assemblage token representation. In a bound-tokens representation, in contrast, grouping is merely a result of mapping the atomic tokens onto a common predicate. At no time during the construction of a bound-tokens representation are the individuals referred to deliberately grouped. Quite to the contrary, the respective atomic tokens have to be kept distinct and are handled separately when mapping each one of them onto the common predicate. Thus, the bound-tokens view does not predict commonality effects that are based on grouping. The idea of bound tokens can therefore not supersede the notion of assemblage tokens suggested in the present paper.

However, bound-tokens representations could be taken into consideration as an alternative to atomic-tokens representations. The main difference between the two types of representations concerns the representation of the predicate. For predicates involving indefinite internal arguments the difference may be crucial. Take for example the sentence both baked a

cake. In an atomic-tokens representation of the kind we discussed in this paper, each of the two individuals referred to by both are separately represented as baking a cake, so that the entire representation contains two cake tokens altogether. Yet, the same does not necessarily hold for a bound-tokens representation. It is conceivable that the whole predicate ‘to bake a cake’ is represented as one complex unit. In this case, the entire representation contains only one cake token rather than two distinct cake tokens. Future research has to show which construal is more appropriate.

Still other kinds of tokens?

The present paper has focused on the mental representations of the referents of plural NPs. The question arises whether in addition to assemblage tokens and atomic tokens still other kinds of tokens have to be envisaged. With regard to this question, collective terms seem especially interesting (e.g., the family, the team, the committee, the choir). One might suppose that the denotation of a singular collective term is mentally represented by an atomic token, as are the denotations of other singular terms (e.g., the car). Yet, entities introduced by a collective term can be re-referred to by a plural anaphor and this is even more natural than using a singular anaphor, at least in English (Bock et al., 1999; Gernsbacher, 1991; Oakhill, Garnham, Gernsbacher, & Cain, 1992) and Spanish (Carreiras & Gernsbacher, 1992). In British English singular collective terms even trigger plural verb agreement under certain conditions (Barker, 1992; for American English see Bock & Eberhard, 1993; Bock et al., 1999). Moreover, collective terms can be combined with certain collective predicates, as (26) illustrates.

(26) The choir assembled in front of the organ.

In semantics, collective terms are therefore often classed with (non-partitioning) plural terms (e.g., Landman, 1989; Link, 1991; but see Barker, 1992). On the other hand collective terms cannot be combined with every collective predicate. For instance, sentences (27) and (28) sound odd, at least if the predicate is meant to express relationships among the members of the group.

(27) ?The choir was dressed differently.

(28) ?The team greeted one another.

Moreover, whereas the assemblage reading of plural NPs can be abolished by inserting an each (cf. (29)), this is not possible with collective terms (cf. (30)).⁶

(29) The men each brought a gift.

(30) ?The team each brought a gift.

These observations suggest that the kind of token representing the referent of a collective term is tighter than an assemblage token (see also Bloom, 1996). This may have something to do with collection-specific properties. Barker (1992) pointed out that the denotation of a collective term has properties of its own, aside from the properties it may inherit from its members. For example, even if the committee denoted in (31) is made up from the exact same men denoted in (32), (32) may be false when (31) is true (the men may have met before because they were old friends, for example).

(31) The committee first met this year.

(32) The men first met this year.

In any case, it seems likely that for collective terms another kind of token has to be postulated, that in contrast to an atomic token has an internal structure, and the individual components of which are more strongly connected to each other than is the case with an assemblage token.

In our view, the different kinds of tokens used for representing discourse referents reflect the way humans structure the world. Although there may be a preferred way of structuring the world, there is certainly some flexibility. Research on lexical specificity and concepts has intensively investigated the flexibility of categorizing an entity at different levels of abstraction (subordinate level, basic level, super-ordinate level). Our study underlines that structuring is also flexible in another respect, concerning grouping and subdividing entities.

References

- Albrecht, J. E., & Clifton, C. (1998). Accessing singular antecedents in conjoined phrases. Memory & Cognition, 26, 599-610.
- Barker, C. (1992). Group terms in English: Representing groups as atoms. Journal of Semantics, 9, 69-93.
- Barsalou, L. W. (1983). Ad hoc categories. Memory & Cognition, 11, 211-227.
- Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural languages. Linguistics and Philosophy, 4, 159-219.
- Bloom, P. (1996). Possible individuals in language and cognition. Current Directions in Psychological Science, 5, 90-94.
- Bloom, P., & Kelemen, D. (1995). Syntactic cues in the acquisition of collective nouns. Cognition, 56, 1-30.
- Bock, K. J., & Eberhard, K. M. (1993). Meaning, sound, and syntax in English number agreement. Language and Cognitive Processes, 8, 57-99.
- Bock, K., Nicol, J., & Cutting, J. C. (1999). The ties that bind: Creating number agreement in speech. Journal of Memory and Language, 40, 330-346.
- Bower, G. H. (1989). Mental models in text understanding. In A. F. Bennett & K. M. McConkey (Eds.), Cognition in individual and social contexts (pp. 129-144). Amsterdam: North Holland.
- Brooks, P. J., & Braine, M. D. S. (1996). What do children know about the universal quantifiers all and each? Cognition, 60, 235-268.
- Carreiras, M. (1997). Plural pronouns and the representation of their antecedents. European Journal of Cognitive Psychology, 9, 53-87.
- Carreiras, M., & Gernsbacher, M. A. (1992). Comprehending conceptual anaphors in Spanish. Language and Cognitive Processes, 7, 281-299.

- Clark, H. H. (1991). Words, the world, and their possibilities. In G. R. Lockhead & J. R. Pomerantz (Eds.), The perception of structure (pp. 263-277). Washington, DC: American Psychological Association.
- van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.
- Dowty, D. (1987). Collective predicates, distributive predicates, and all. Proceedings of the Eastern States Conference on Linguistics, ESCOL3.
- Eschenbach, C., Habel, C., Herweg, M., & Rehkämper, K. (1989). Remarks on plural anaphora. Proceedings of the fourth conference of the European Chapter of the Association for Computational Linguistics (pp. 161-167). Manchester.
- Ferreira, F., & Clifton, C. Jr. (1986). The independence of syntactic processing. Journal of Memory and Language, 25, 348-368.
- Franklin, N., Tversky, B., & Coon, V. (1992). Switching points of view in spatial mental models. Memory & Cognition, 20, 507-518.
- Fraurud, K. (1991). What do plural pronouns refer to? (GAP-AP20). Hamburg: University of Hamburg, Department of Informatics.
- Frazier, L., Pacht, J. M., & Rayner, K. (1999). Taking on semantic commitments, II: collective versus distributive readings. Cognition, 70, 87-104.
- Garnham, A. (1996). The other side of mental models: theories of language comprehension. In J. Oakhill & A. Garnham (Eds.), Mental models in cognitive science (pp. 35-52). Hove: Psychology Press.
- Garrod, S. C., & Sanford, A. J. (1982). The mental representation of discourse in a focussed memory system: implications for the interpretation of anaphoric noun phrases. Journal of Semantics, 1, 21-41.

- Garrod, S. C., & Sanford, A. J. (1998). The role of scenario mapping in text comprehension. Discourse Processes, 26, 159-190.
- Gelman, S. A., & Tardif, T. (1998). A cross-linguistic comparison of generic noun phrases in English and Mandarin. Cognition, 66, 215-248.
- Gernsbacher, M. A. (1990). Language comprehension as structure building. Hillsdale, NJ: Erlbaum.
- Gernsbacher, M. A. (1991). Comprehending conceptual anaphors. Language and Cognitive Processes, 6, 81-105.
- Gernsbacher, M. A. (1997). Two decades of structure building. Discourse Processes, 23, 265-304.
- Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. Journal of Memory and Language, 31, 129-151.
- Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. Journal of Memory and Language, 26, 69-83.
- Grosz, B., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. Computational Linguistics, 12, 175-204.
- Herweg, M. (1988). Ansätze zu einer semantischen und pragmatischen Theorie der Interpretation pluraler Anaphern (GAP-AP 2). Hamburg: University of Hamburg, Department of Informatics.
- Hielscher, M., & Müsseler, J. (1990). Anaphoric resolution of singular and plural pronouns: The reference to persons being introduced by different co-ordinating structures. Journal of Semantics, 7, 347-364.
- Hörmann, H. (1983). The calculating listener, or: how many are 'einige', 'mehrere' and 'ein paar' ('some', 'several', and 'a few'). In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), Meaning, use, and interpretation of language (pp. 221–234). Berlin: de Gruyter.

- Jackendoff, R. (1983). Semantics and cognition. Cambridge, MA: MIT Press.
- Jackendoff, R. (1987). Consciousness and the computational mind. Cambridge, MA: MIT Press.
- Johnson-Laird, P. N. (1983). Mental models. Cambridge: Cambridge University Press.
- Kamp, H., & Reyle, U. (1993). From discourse to logic. Dordrecht: Kluwer Academic Publishers.
- Landman, F. (1989). Groups. Linguistics and Philosophy, 12, 559-605 (part I) and 723-744 (part II).
- Lang, E. (1984). The semantics of coordination. Amsterdam: Benjamins.
- Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), Meaning, use, and interpretation of language (pp. 302-323). Berlin: de Gruyter
- Link, G. (1991). Plural. In A. von Stechow & D. Wunderlich (Eds.), Semantik (pp. 418-440). Berlin: de Gruyter.
- Massey, G. J. (1976). Tom, Dick, and Harry, and all the king's men. American Philosophical Quarterly, 13, 89-107.
- Markman, E. M. (1990). Constraints children place on word meanings. Cognitive Science, 14, 57-77.
- Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? Annual Review of Psychology, 51, 121-147.
- Morrow, D. (1994). Spatial models created from text. In H. van Oostendorp & R. A. Zwaan (Eds.), Naturalistic text comprehension (pp. 57-78). Norwood, NJ: Ablex Publishing.
- Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. Journal of Memory and Language, 28, 292-312.

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1990). Situation-based inferences during narrative comprehension. In A. C. Graesser & G. H. Bower (Eds.), The psychology of learning and motivation: Vol. 25. Inferences and text comprehension (pp. 123-135). San Diego: Academic Press.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. Journal of Memory and Language, 26, 165-187.

Murphy, G. L. (1984). Establishing and accessing referents in discourse. Memory & Cognition, 12, 489-497.

Müsseler, J., Hielscher, M., & Rickheit, G. (1995). Focussing in spatial mental models. In G. Rickheit & C. Habel (Eds.), Focus and coherence in discourse processing (pp. 35-52). Berlin: de Gruyter.

Oakhill, J., Garnham, A., Gernsbacher, M. A., & Cain, K. (1992). How natural are conceptual anaphors? Language and Cognitive Processes, 7, 257-280.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: a suggestion for a better and more powerful analysis. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 785-794.

Reed, A. M. (1991). On interpreting partitives. In D. J. Napoli & J. A. Kegl (Eds.), Bridges between psychology and linguistics. A Swarthmore Festschrift for Lila Gleitman (pp. 207-223). Hillsdale, NJ: Lawrence Erlbaum.

Rehkämper, K. (1991). Zur Behandlung pluraler Pronomen in der Theorie des Textverstehens von Sanford & Garrod. Sprache & Kognition, 10, 136-147.

Reis, M., & Vater, H. (1980). Beide. In G. Brettschneider & C. Lehmann (Eds.), Wege zur Universalienforschung. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler (pp. 365-391). Tübingen: Narr.

Rinck, M., & Bower, G. H. (1995). Anaphora resolution and the focus of attention in situation models. Journal of Memory and Language, 34, 110-131.

Sanford, A. J., & Lockhart, F. (1990). Description types and method of conjoining as factors influencing plural anaphora: A continuation study of focus. Journal of Semantics, 7, 365-378.

Sanford, A. J., & Moxey, L. M. (1995). Notes on plural reference and the scenario-mapping principle in comprehension. In G. Rickheit & C. Habel (Eds.), Focus and coherence in discourse processing (pp. 18-34). Berlin: de Gruyter.

Scha, R. (1981). Distributive, collective and cumulative quantification. In J. Groenendijk, T.M.V. Janssen, & M. Stokhof (Eds.), Formal methods in the study of language, Part 2 (pp. 483-512). Amsterdam: Mathematisch Centrum.

Schopp, A. (1995). Focussing and the use of German beide. In G. Rickheit & C. Habel (Eds.), Focus and coherence in discourse processing (pp. 75-93). Berlin: de Gruyter.

Schwarzschild, R. (1994). Plurals, presuppositions and the sources of distributivity. Natural Language Semantics, 2, 201-248.

Shipley, E. F., & Shepperson, B. (1990). Countable entities: developmental changes. Cognition, 34, 109-136.

Verkuyl, H., & van der Does, J. (1996). The semantics of plural noun phrases. In J. van der Does & J. van Eijck (Eds.), Quantifiers, logic, and language (pp.337-375). Cambridge: Cambridge University Press.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. Cognition, 61, 261-298.

Wilson, G., & Stevenson, R. J. (1997). Spatial information and noun phrase conjunction as cues to group formation. In B. Kokinov (Ed.), Perspectives in cognitive science. Vol. IV. Sofia: New Bulgarian University Press.

Winter, Y. (1998). Atom predicates and set predicates: towards a general theory of plural quantification. Proceedings of Semantics and Linguistic Theory (SALT) 8.

Winter, Y. (2000). Distributivity and dependency. Natural Language Semantics, 8, 27-69.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. Psychological Bulletin, 123, 162-185.

Appendix

Materials Used in Experiment 1

Stimulus sentences were in German. The English translations given here are as near as possible to a word-by-word translation. For the German sich, however, which is a pronoun coindexed with the grammatical subject, no translation is given (items #4 - #7, #13 - #15). In our proposal, sich when coindexed with a non-partitioning plural expression like sie, indicates a reflexive relation at the coarse granularity level. The assemblage denoted by the subject NP is asserted to be related to itself, leaving open whether each atomic individual is related to itself or to other atomic individuals of the assemblage. In contrast, when coindexed with a partitioning expression such as beide, the reflexive sich unequivocally specifies that each atomic individual is related to itself. Thus, the German sich is neither equivalent to each other nor to themselves.

Unconstrained sentence-judgment task (a. stimulus sentence, b. question)

- 1a. Sie / Beide kauften fünf Tafeln Schokolade.
They / Both bought five chocolate bars.
- 1b. Wie viele Tafeln Schokolade wurden gekauft?
How many chocolate bars were bought?
- 2a. Sie / Beide bestellten ein Taxi.
They / Both called a taxi.
- 2b. Wie viele Taxen wurden bestellt?
How many taxis were called?
- 3a. Sie / Beide brachten ein Geschenk mit.
They / Both brought a gift.
- 3b. Wie viele Geschenke wurden mitgebracht?
How many gifts were brought?

- 4a. Sie / Beide fanden sich attraktiv.
They / Both thought [pronoun coindexed with subject] attractive.
- 4b. Wer fand wen attraktiv?
Who thought whom attractive?
- 5a. Sie / Beide klopfen sich auf die Schulter
They / Both patted [pronoun coindexed with subject] on the back.
- 5b. Wer klopfte wem auf die Schulter?
Who patted whom on the back?
- 6a. Sie / Beide beschuldigten sich der Unaufmerksamkeit.
They / Both accused [pronoun coindexed with subject] of inattentiveness.
- 6b. Wer beschuldigte wen?
Who accused whom?
- 7a. Sie / Beide heirateten im August.
They / Both married in August.
- 7b. Wieviele verschiedene Hochzeiten fanden statt?
How many different weddings took place?
- 8a. Sie / Beide telefonierten eine Stunde lang.
They / Both telephoned for an hour.
- 8b. Wie viele verschiedene Telefongespräche fanden statt?
How many different telephone conversations took place?
- 9a. Am Donnerstag spielten sie / beide Tennis.
On Thursday, they / both played tennis.
- 9b. Spielten Fritz und Maria gemeinsam Tennis?
Did Fritz and Maria play tennis together?

Constrained sentence-judgment task.

10. Sie / Beide trugen saubere Hemden.
They / Both wore clean shirts.
11. Sie / Beide erhielten die Diplomzeugnisse aus der Hand des Rektors.
They / Both received the diploma certificates from the rector.
12. Sie / Beide gingen in die Duschkabinen.
They / Both went into the shower cabins.
13. Sie / Beide küssten sich auf dem Sofa.
They / Both kissed [pronoun coindexed with subject].
14. Sie / Beide unterschieden sich in der Haarfarbe.
They / Both differed [pronoun coindexed with subject] in hair color.
15. Sie / Beide trafen sich am Kino.
They / Both met [pronoun coindexed with subject] at the cinema.
16. Sie / Beide waren gleichen Alters.
They / Both were of the same age.
17. Sie / Beide waren gleich groß.
They / Both were equally tall.
18. Sie / Beide waren unterschiedlicher Nationalität.
They / Both were of differing nationality.

Author Note

Barbara Kaup, Psychology Department; Stephanie Kelter, Department of Psychology; Christopher Habel, Department of Informatics.

The research reported in this paper was partly supported by a grant from the Deutsche Forschungsgemeinschaft (Ha 1237/5-2) to Christopher Habel and Stephanie Kelter. We would like to thank Berry Claus, Carola Eschenbach, Ben Hofer, Annette Lessmöllmann, Christie Manning, Christian Stein, and Rolf Zwaan, who provided valuable comments on earlier versions of the manuscript, and Ute Schmid, who discussed our questions with American native speakers.

Correspondence concerning this article should be addressed to B. Kaup, Psychology Department of Florida State University, Tallahassee, Florida 32306-1270, USA. Electronic mail may be sent to kaup@psy.fsu.edu.

Footnotes

¹Note, the term predicate refers to the property, state, or action that is expressed by the (whole) verb phrase of a sentence. For a sentence such as The girl is knitting a sock, the predicate is 'knitting a sock', not just 'knitting'.

²In contrast, the universal quantifier all does not generally promote an individual-tokens representation. It does so when it is a floating quantifier (e.g., The men all brought a birthday present) and when it is the determiner of a partitive NP (All of the pupils brought a birthday present). However, a simple definite NP such as all the children seems to trigger either an atomic-tokens representation or an assemblage-token representation, with no clear bias towards one or the other (Brooks & Braine, 1996 Exp. 3). Moreover, such a subject NP can be combined with a collective predicate (All the pupils gathered at the station). Because of this ambiguity, we will ignore this type of NP in what follows.

³Here and in the following the terms "anaphoric expression" and "anaphor" are meant to denote discourse anaphors, i.e. anaphors crossing the sentence boundaries. Moreover, anaphors are not analyzed in terms of relationships between anaphoric expressions and preceding linguistic expressions (i.e., antecedents), but in terms of relationships between (mental representations of) linguistic expressions and (mental representations of) non-linguistic entities (i.e. the referents).

⁴At the end of the questionnaire there were four additional items designed to explore the corresponding hypotheses with respect to language production. However, because of their exploratory nature, we will not discuss the design and results of these items in the present article.

⁵ This claim implies, among other things, that a collective predicate does not have the potential to trigger an assemblage interpretation of the subject NP. Presupposing that the applicability of a collective predicate is strictly confined to assemblages, this leads to the

hypothesis that the combination of a collective predicate and a partitioning subject NP yields an unacceptable sentence (e.g., both orphans differed in hair color.). The results of Experiment 1 (constrained task) provided positive evidence for this hypothesis. It should be noted, however, that there seem to be specific linguistic devices for overriding the atomic interpretation suggested by a partitioning NP, at least if the predicate is of the unconstrained type. For example, inserting together in a sentence with a partitioning NP and an unconstrained predicate (e.g., Both orphans together brought a gift) seems to induce an assemblage interpretation of the subject NP (the sentence is interpreted as describing a situation in which one gift was brought, not two). There are also linguistic means (e.g., a floating each or both) that have the converse effect, namely that the denotation of a non-partitioning NP is conceived of as a set of distinct individuals rather than as an assemblage (the orphans brought a gift vs. the orphans each brought a gift). Thus, strictly speaking, the claim that non-partitioning expressions routinely lead to assemblage-token representations and partitioning expressions to atomic-token representations, only holds for sentences that do not contain any of these specific linguistic devices.

⁶We thank Christopher M. Hogan, who brought this issue to our attention.

Table 1. Mean Acceptability Ratings for Sentences of the Unconstrained Set in Experiment 1, as a Function of Pronoun in Stimulus Sentence and Category of Interpretation.

Category of Interpretation		sie (they)		beide (both)	
		Acceptabilit y	n ^a	Acceptabilit y	n ^a
1	Assemblage Only	1.41	212	2.12	79
2	Assemblage Preferred	2.04	21	2.07	14
3	No Preference	2.57	9	2.93	9
4	Atomic Preferred	2.67	7	1.98	15
5	Atomic Only	3.18	11	1.97	147
	Mean (Unweighted)	2.38		2.21	

^a total number of responses in the respective cell. Because of the blanks (see Scoring), the numbers of responses to they- and both-sentences do not each add to 270, but to 260 and 264, respectively.

Table 2

Sample Story of Experiment 2 with its Four Commonality Versions (Translated From the German Original)

Setting	Mary is doing the dishes in the kitchen. She wants the apartment to look as nice as possible for the house warming party that will be held in the evening.
Different-Far	Harry cannot help with drying the dishes because he promised his colleagues that he would bring some cookies to tomorrow's meeting at work. He is currently on his way to the bakery to buy the cookies.
Same-Far	Harry is currently buying some fresh cookies for the party at the bakery. On the way, he also does some shopping for tomorrow's meeting at work.
Different-Close	Harry cannot help with drying the dishes because he promised his colleagues to bake some cookies for tomorrow's meeting at work. He is currently taking the cookies out of the oven.
Same-Close	Harry made some cookies for the party and is currently waiting for them to get finished so that he can take them out of the oven. While he is waiting, he dries the dishes.
Target	They/Both are enjoying the new apartment.
Post-Target	They are curious if the guests will like the apartment, too.

Note. Each participant was presented with only one of the versions.

Figure Captions

Figure 1. The atomic-tokens view. Discourse-level representation for the sentence The orphans laughed.

Figure 2. The assemblage-token view. Discourse-level representation for the sentence The orphans laughed.

Figure 3. The assemblage-token view. Discourse-level representation for the sentence The acrobats rode into the ring on white horses.

Figure 4. Mean residuals of reading times for the target sentence and the post-target sentence in the they-they condition and the both-they condition of Experiment 2.

Figure 5. The bound-tokens view. Discourse-level representation of the sentence The orphans laughed.