Between Intersective and ‘Split’ Interpretations of Predicate Conjunction: The Role of Typicality

Eva B. Poortman (e.b.poortman@uu.nl)
Utrecht Institute of Linguistics OTS, Utrecht University
Trans 10, 3512 JK, Utrecht, The Netherlands

Abstract

This paper aims to account for the different interpretations of plural sentences with conjoined predicates. Such sentences are sometimes interpreted strictly intersectively, sometimes strictly non-intersectively (or ‘split’), and sometimes both interpretations appear to be allowed. We propose that the logical interpretation of these sentences is systematically affected by typicality effects of the complex predicate in the sentence. We show with a set of experiments that a) the acceptability of a non-intersective interpretation can be expressed in terms of a continuum and b) each acceptability proportion is predicted by the typicality of the two conjoined predicates applying simultaneously. This way, we specify at least one of the relevant pragmatic considerations that determine the interpretation of a plural sentence with conjunctive predicates. Moreover, these results stress the importance of conceptual structure of predicates for compositional semantics.

Keywords: conjunction; plurality; plural predication; typicality effects; concepts; reasoning.

Introduction

Plural sentences with conjunctive predicates as in (1) and (2) are considered to be true if and only if every boy that is referred to is in the intersection of the two sets that are denoted by the conjoined verbs. In other words, sentence (1) is true iff each boy is both sitting and reading, and sentence (2) is true iff each boy is both waving and smiling.

(1) The boys are sitting and reading
(2) The boys are waving and smiling

We arrive at such interpretations by applying the well-known boolean analysis of conjunction, according to which it behaves as set-theoretic intersection (Keenan & Faltz, 1985; Partee & Rooth, 1983), and combining it with a distributivity operator (Link, 1983). Importantly, such an analysis assumes that the way we reason about these natural language sentences is independent of lexical information – thus assuming a clear division between compositional and lexical semantics. As a result, the logical interpretations of sentences like (3) and (4) are expected to be derived in a similar way as those of (1) and (2), with the difference between the sentences only being a matter of word meaning.

(3) The boys are sitting and standing
(4) The boys are sitting and cooking

In this paper, we report experiments that show that sentences (3) and (4) in fact receive weaker logical interpretations than sentences (1) and (2). Sentence (3) is generally interpreted such that a subset of the boys is sitting and the rest of the boys are standing – we will call this a ‘split’ interpretation (Heycock & Zamparelli, 2005). Sentence (4) also allows such a ‘split’ interpretation, but crucially to a lesser extent than sentence (3). Understanding such acceptability patterns calls for a systematic investigation of the lexical information in the sentence, as this appears to be inseparable from a proper analysis of conjunction. We show that there is in fact a continuum of acceptability values for non-intersective or ‘split’ interpretations of sentences with conjunctive predicates, and we account for this continuum with a principle that predicts how language users apply predicates to plural subjects based on typicality structure of the complex predicate.

Many previous works have described non-intersective interpretations of plural sentences (e.g. Kriška, 1990; Heycock & Zamparelli, 2005; Winter, 2001). Winter (2001) tries to systematically account for the different interpretations, in which properties are distributed differently over individual entities. He does so by extending the Strongest Meaning Hypothesis (SMH) that was put forward by Dalrymple et al. (1998). The SMH aims to resolve ambiguity that is caused by lexical and other contextual information, specifically in the area of reciprocals. Dalrymple et al. claim that the varying logical interpretations of structurally similar reciprocal sentences as in (5) and (6) are captured by their principle.

(5) The boys know each other
(6) The boys are following each other

For each occurrence of the reciprocal, this principle selects as its interpretation the strongest meaning (from an inventory of six possible meanings) that is consistent with context. For example, if we assume three boys, then sentence (5) receives a strong interpretation in which every boy knows every other boy, since there are no contextual restrictions on the amount of possible “knowing-relations”. All weaker meanings are consequently disallowed for (5). By contrast, sentence (6) most likely means something weaker than every boy following every other boy.

\(^3\) Throughout the paper, when using the term “interpretation” we simply refer to the situations that support a truthful usage of a sentence.
According to the SMH, we weaken the meaning of the sentence as far as context pushes us to. For this example, that meaning is most likely one where boy 1 follows boy 2, and boy 2 follows boy 3. This is the strongest candidate meaning that does not contradict our knowledge about following people.

Winter (2001) proposes that a maximality principle like the SMH is not construction-specific to plural sentences with reciprocals. He rephrases the SMH into a general principle of plural predication, such that any complex plural predicate with a meaning that is derived from one or more singular predicates using universal quantification is interpreted using the logically strongest truth conditions that are not contradicted by known properties of the singular predicate(s) (Winter, 2001). The contrast between minimal pairs like (1) and (3) is then captured in the following way. Again, the SMH selects the logically strongest possible candidate meaning for each sentence. When a strong interpretation (intersective conjunction) is consistent with properties of the predicates, then this is the attested meaning of the sentence – an example is sentence (1). On the other hand, when such a strong interpretation is inconsistent with these properties, the interpretation is weakened. We see this in sentence (3): An intersective interpretation in which all boys are in the intersection of the set of sitting individuals and the set of standing individuals contradicts what we know about ‘sitting’ and ‘standing’. Thus, sentence (3) receives a ‘split’ interpretation, which is the strongest interpretation that does not contradict this knowledge.

In the current paper, we argue that the predictions made by the SMH can be too strong. Consider sentences (7) and (8), which are of a similar nature to sentence (4) above.

(7) The men are lying down and drinking
(8) The men are waving and drawing

If Winter's extended SMH is correct in assuming that non-intersective interpretations are only available when intersective interpretations are strictly ruled out by the predicates, then these sentences would only allow an intersective interpretation. As we will show in the current paper, non_intersective, ‘split’ interpretations are readily available to many speakers for sentences like (4), (7) and (8), even though the predicates do not strictly exclude an intersective one. For example, it may be exceptional, but it is possible for a person to sit and cook simultaneously.

Several previous works have recognized a similar problem for the SMH concerning reciprocal sentences that receive weaker interpretations than predicted (e.g. Winter 2001; Philip, 2000; Kerem, Friedmann & Winter, 2009; Struiksma et al., submitted). Struiksma et al. (submitted), for example, showed that a sentence like The boys are pinching each other in the case of three boys is judged as true in a situation where each boy pinches only one other boy, despite the fact that a stronger interpretation is not excluded by properties of the predicate pinch.

These examples point to a fundamental issue with the proposal at hand. The SMH, both in its original and extended form, assumes that the interpretation of these sentences is only sensitive to so-called “definitional” aspects of the meaning of predicate concepts. In other words, it only takes into account whether particular denotations of predicates are possible or impossible, i.e. whether they are an instance of that predicate concept or not. In the case of predicate conjunction, that means that the hypothesis only looks at whether intersective conjunction is possible or not, given the predicates at hand. Such sharp distinctions appear to be insufficient in accounting for the interpretation patterns that we observe. Alternatively, one can take into account typicality effects in categorization. The notion of typicality refers to the phenomenon that human subjects are able to grade different instances of a concept with respect to their representativeness of a given category. Since the 1970’s, a range of psychological studies has shown for one-place predicates that subjects consistently rank some instances of a concept as more typical than others, and that such rankings correlate with other measures of typicality such as categorization speed and error rate (e.g. Rosch, 1973; Smith, Shoben & Rips, 1974; Rosch & Mervis, 1975).

Taking into account typicality effects on reasoning was first proposed as a solution for reciprocal sentences, in the shape of the Maximal Typicality Hypothesis (Kerem et al., 2009; Struiksma et al., submitted). This hypothesis assumes that typicality effects also exist for binary predicate concepts (like pinch), and that these systematically affect the logical interpretation of the reciprocal expression that they combine with. In this paper, we extend the same logic to plural sentences with predicate conjunction. We claim that typicality effects also affect interpretation in plural sentences where two predicate concepts are conjoined. The proposal works as follows. For a complex concept like lying down and drinking, a situation where one is both lying down and drinking at the same time is an instance of that concept, but most likely an atypical one, at least to a certain extent. Consequently, when such a complex predicate combines with a plural, this affects the degree to which language users diverge from an intersective interpretation. Crucially, we predict that both the notions of typicality and acceptability can be expressed in terms of a continuum – allowing for more subtle distinctions than the SMH.

To summarize, our proposal predicts the following: a) that there is a continuum of typicality effects for complex predicates made up of two conjoined singular predicates, b) that there is a continuum of acceptability values for a particular interpretation of a plural sentence with those complex predicates and c) that the values on both continuums correlate – indicating that typicality for concepts in isolation systematically affects interpretation of sentences containing those concepts. We conducted two behavioral experiments and a correlation analysis to test these predictions.
Experimental Investigation

This section reports on pretests, two experiments and a correlation analysis. Experiment 1 checked the acceptability of plural predicate conjunction sentences of the form The A are X and Y (where A is a plural noun and X and Y are predicates) in a non-intersective, ‘split’ situation. Experiment 2 measured typicality effects for complex predicate concepts in isolation. Materials for the experiments were constructed based on pretests that were conducted in order to include a wide range of typicality values in the actual experiments.

Pretests: Constructing Materials

The aim of the first pretest was to very informally gather as many Dutch verb combinations X and Y as possible, especially atypical ones. We provided 8 subjects with sets of two pairs in which X was identical: one very natural pair, and one pair that is physically impossible to apply simultaneously, e.g. sitting and reading and sitting and standing. We then asked them to provide as many verbs that they could come up with that combine with X (i.e. sitting in this case) that led to a possible but atypical, uncommon or strange combination. These pairs, combined with more natural pairs that we came up with, led to a list of 91 verb combinations in total.

In the second pretest, we had 29 different subjects rate these 91 pairs for compatibility. For each pair, subjects were asked to rate how odd they would consider it if both verbs applied to one person at the same time. Oddness was rated on a 6-point scale, where 1 meant “not odd at all” and 6 meant “physically impossible”. We mentioned explicitly that 5 thus meant “very odd, but physically possible”, in order to distinguish large atypicality from impossibility. Results of this pretest showed great variability in ratings between verb pairs. We first selected 12 sets of verb pairs with identical X that showed greatest variability within the set. From each of these 12 sets, we selected three verb pairs for the experiments: the verb pair that was rated lowest on the oddness scale (compatible pairs like sitting and reading), the verb pair that was rated highest (incompatible ones like sitting and standing), and a verb pair that was rated in between, at a mean of 4 points (atypical pairs like sitting and cooking). An overview of the 36 verb pairs that constituted the final material (translated from Dutch) is given in table 1. Creating the three groups (with labels ‘compatible’, ‘incompatible’ and ‘atypical’) was done purely to ensure variability, and for the sake of clarity we will refer to the three groups when discussing set-up and results of experiments 1 and 2. Note however that they are not meaningful in the final correlation analysis of all data points.

Experiment 1: Interpretation of Plural Predicate Conjunction Sentences

This experiment checked the acceptability of 36 plural sentences with two conjoined verbs in a ‘split’ situation. Each sentence was of the form The A are X and Y (where A is a plural noun and X and Y are verbal predicates).

Participants A total of 33 students from Utrecht University (28 female, age M = 21) participated for monetary compensation. All participants were native speakers of Dutch without dyslexia. Prior to the experiment all participants signed an informed consent form.

Materials The material consisted of two versions of a truth-value judgment task, each containing 18 unique test items plus 18 filler items that were the same across versions. Each test item contained a plural predicate conjunction sentence in Dutch (The A are X and Y) and a drawing depicting four individuals in a non-intersective, ‘split’ interpretation of that sentence: predicate X applied only to persons 1 and 2, predicate Y applied only to persons 3 and 4. Half of the pictures depicted male individuals, and the other half depicted female individuals. An example of a test item drawing is given in figure 1.

Figure 1: Example of a test item drawing of experiment 1.

In each version of the experiment, one third of the test items contained sentences with verb pairs that were considered compatible X and Y in the second pretest (e.g. The men are sitting and reading), one third contained sentences with verb pairs that were considered incompatible X and Y (e.g. The men are sitting and standing) and one third contained sentences with pairs that were considered atypical X and Y (e.g. The men are sitting and cooking). The same drawings

---

2 Phrasing the question negatively by asking “how odd” subjects would rate a situation was done because directly asking for “how typical” they would judge a situation turned out to be ambiguous in Dutch. Some subjects interpreted the word typical to mean “atypical”, whereas asking for oddness is unambiguous.

3 Additional inclusion criteria included that each verb should be expressed by one word only, ratings for verb pairs should have smallest variation, and atypical verb pairs should have no “6” point ratings.

4 We measured acceptability of the sentence given a ‘split’ situation because sentences with incompatible pairs cannot be depicted any other way, and we wished to keep all factors in the comparison between pairs equal.
Table 1: Overview of predicate pairs, translated from Dutch (the light and dark cells represent the division of items over the two versions of experiment 1).

<table>
<thead>
<tr>
<th>Compatible</th>
<th>Incompatible</th>
<th>Atypical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting and reading</td>
<td>Sitting and standing</td>
<td>Sitting and cooking</td>
</tr>
<tr>
<td>Waving and smiling</td>
<td>Waving and clapping</td>
<td>Waving and drawing</td>
</tr>
<tr>
<td>Walking and singing</td>
<td>Walking and swimming</td>
<td>Walking and writing</td>
</tr>
<tr>
<td>Crawling and screaming</td>
<td>Crawling and jumping</td>
<td>Crawling and reading</td>
</tr>
<tr>
<td>Standing and reading</td>
<td>Standing and squatting</td>
<td>Standing and falling asleep</td>
</tr>
<tr>
<td>Reading and smiling</td>
<td>Reading and sleeping</td>
<td>Reading and drawing</td>
</tr>
<tr>
<td>Lying down and stretching</td>
<td>Lying down and running</td>
<td>Lying down and drinking</td>
</tr>
<tr>
<td>Drawing and yawning</td>
<td>Drawing and typing</td>
<td>Drawing and walking</td>
</tr>
<tr>
<td>Swimming and smiling</td>
<td>Swimming and crawling</td>
<td>Swimming and reading</td>
</tr>
<tr>
<td>Texting and frowning</td>
<td>Texting and knitting</td>
<td>Texting and waving</td>
</tr>
<tr>
<td>Knitting and singing</td>
<td>Knitting and clapping</td>
<td>Knitting and walking</td>
</tr>
<tr>
<td>Sleeping and drooling</td>
<td>Sleeping and telephoning</td>
<td>Sleeping and standing</td>
</tr>
</tbody>
</table>

were used for sentences with compatible and incompatible pairs with identical X (e.g. The men are sitting and standing and The men are sitting and reading). To ensure that subjects never saw the same drawing twice (such as the one in figure 1), one of these sentences occurred in version 1 and the other occurred in version 2. The atypical items were divided over the two versions, resulting in two experiments with 6 sentences with compatible pairs, 6 sentences with incompatible pairs and 6 sentences with atypical pairs each, accompanied by 18 unique drawings. The items of each version are represented by light and dark cells in table 1.

Filler items contained similar drawings with four people, but a different type of accompanying sentence. The accompanying sentences in the filler items were either sentences with quantifiers (Some boys are X) or sentences mentioning specific individuals in the picture (Boys A, B and C are X). Half of the filler items were expected to be judged true, and half of them were expected to be judged false. Both versions of the experiment contained the same filler items.

The order of items was pseudo-randomized using Mix software (Van Casteren & Davis, 2006), with the following restrictions: items containing the same verb were at least six items apart; there were at most two test items immediately following each other, and at most two filler items immediately following each other; similar test items (in terms of compatible/incompatible/atypical) or similar filler items (in terms of quantifier/specific individuals) never immediately followed each other. Finally, we constructed two orders of each version, with the second one having reversed order of items.

Procedure Each participant completed one version of the experiment. The task was presented in a sound-proof booth on a PC using Open Sesame software (Mathôt, Schreij & Theeuwes, 2012). Prior to entering the sound-proof booth, each participant received verbal instructions explaining the experimental set-up. Further, more detailed instructions were given on the PC monitor.

After being instructed, each participant completed three practice trials. Subsequently, they were given the opportunity to ask for clarifications, if necessary. All verbs used in the practice session did not appear in the actual experiment. The experiment itself consisted of the 36 items described above. Drawing and sentence were presented in the center of a white screen. Participants were instructed to indicate as soon as possible whether they judged the sentence to be true or false given the situation in the drawing by pressing the left or right button with their dominant hand.

Coding and analysis Responses were coded “1” when participants judged a sentence to be true for a given drawing, and “0” when they judged a sentence to be false. This way we computed the proportion of acceptability of a sentence for a given drawing. We performed a repeated measures ANOVA with Compatibility as the within-subjects factor (with 3 levels: compatible, atypical, and incompatible)². Post-hoc Bonferroni corrected multiple comparisons were performed in order to analyze differences between different Compatibility levels in detail.

Results Table 2 provides an overview of the data. It shows the mean acceptability of sentences (in proportions) for the three levels of Compatibility that were tested, for all versions taken together. Overall, mean acceptability per verb pair ranged from .24 to 1.

Table 2: Mean acceptability ratings (in proportions).

<table>
<thead>
<tr>
<th>Compatibility type</th>
<th>Mean acceptability (st. dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compatible</td>
<td>.54 (.31)</td>
</tr>
<tr>
<td>Atypical</td>
<td>.78 (.31)</td>
</tr>
<tr>
<td>Incompatible</td>
<td>.84 (.32)</td>
</tr>
</tbody>
</table>

² A repeated measures ANOVA with Version as between-subjects factor was also performed, but showed no effect of Version (F (3, 29) = .47, p = .71). We thus collapsed the versions for the analysis.
A repeated measures ANOVA revealed that there was a main effect of Compatibility \( (F(1.36, 43.49) = 37.41, p < .001) \). This means that the mean proportions of acceptability for the three Compatibility levels are not equal. Pairwise comparisons show that all three levels differ significantly from each other in acceptability: the acceptability of sentences with compatible predicates differs from the acceptability of sentences with incompatible predicates \( (p < .001) \); the acceptability of sentences with compatible predicates differs from the acceptability of sentences with atypical predicates \( (p < .001) \); and the acceptability of sentences with incompatible predicates differs from the acceptability of sentences with atypical predicates \( (p < .05) \).

**Experiment 2: Typicality Effects for Predicate Pairs**

This experiment checked typicality effects for the 36 complex concepts that were used in sentences of experiment 1. We measured the typicality of one particular instance of each complex predicate, namely the one in which both predicates apply simultaneously to one individual.

**Participants** The same 33 students from Utrecht University from experiment 1 partici-pated in this experiment. Each subject completed the interpretation experiment first, before proceeding with the typicality experiment. Also, in between experiments they took part in a third, unrelated experiment.

**Materials** The materials consisted of a questionnaire containing 36 statements about one person involved in two actions simultaneously. Half of the statements were about males and half of them were about females (matching the gender of persons in the pictures of experiment 1). Each statement contained a singular object (a male or a female) and two conjoined predicates (e.g., *The man is sitting and reading*). The 36 pairs of verbs were the same as the ones used in sentences of experiment 1, thus one third of the pairs were considered compatible in the second pretest (e.g., *sitting and reading*), one third were considered incompatible (e.g., *sitting and standing*), and one third were considered atypical (e.g., *sitting and cooking*).

The order of items was pseudo-randomized using Mix software (Van Casteren & Davis, 2006), with the restriction that at most two items of the same type (in terms of compatible/incompatible/atypical) immediately followed each other.

Finally, we constructed four different orders of the questionnaire: two versions that started with the statements about males (with the second one having reversed order within males and females statements), and two versions that started with the statements about females (with the second one having reversed order within males and females statements).

**Procedure** Each participant received one of the questionnaires on paper, in a sound-proof booth. They were instructed to rate how odd⁶ they would consider it if both verbs applied to the given person at the same time. Oddness was rated on a 6-point scale, where 1 meant “not odd at all” and 6 meant “physically impossible”. We mentioned explicitly that 5 thus meant “very odd, but physically possible”, in order to distinguish large atypicality from impossibility.

**Coding and analysis** Responses were coded “1” through “6” corresponding to the participant’s judgment. This way we computed the atypicality rating of two verb pairs applying simultaneously. We performed a repeated measures ANOVA with Compatibility as the within-subjects factor (with 3 levels: compatible, atypical, and incompatible). Post-hoc Bonferroni corrected multiple comparisons were performed in order to analyze differences between different Compatibility levels in detail.

**Results** Table 3 provides an overview of the data. It shows the mean atypicality rating for the three levels of Compatibility that were tested, for all versions taken together. Overall, mean ratings per verb pair ranged from 1.03 to 5.94.

<table>
<thead>
<tr>
<th>Compatibility type</th>
<th>Mean atypicality (st. dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compatible</td>
<td>1.35 (.29)</td>
</tr>
<tr>
<td>Atypical</td>
<td>3.82 (.58)</td>
</tr>
<tr>
<td>Incompatible</td>
<td>5.66 (.38)</td>
</tr>
</tbody>
</table>

A repeated measures ANOVA revealed that there was again a main effect of Compatibility \( (F(1.95, 62.45) = 1187.02, p < .001) \). This means that the mean atypicality ratings for the three Compatibility levels are not equal. Pairwise comparisons show that all three levels differ significantly from each other: the atypicality of supposed compatible pairs applying simultaneously differs from the atypicality of supposed incompatible pairs applying simultaneously \( (p < .001) \); similarly for the atypicality of compatible vs. atypical pairs \( (p < .001) \); and the atypicality of incompatible vs. atypical pairs \( (p < .001) \). This means that the three groups that we selected based on the pretest were confirmed in the typicality experiment (with different subjects and a subset of the stimuli).

**Correlation Between Interpretation and Typicality**

The crucial test for our proposal is the relationship between interpretation and typicality. In order to account for the degree to which non-interactive interpretations of sentences are available given two particular conjoined predicates, we

---

⁶ Again, phrasing the question negatively by asking “how odd” subjects would rate a situation was done because directly asking for “how typical” they would judge a situation turned out to be ambiguous in Dutch. Some subjects interpreted the word *typical* to mean “atypical”, whereas asking for oddness is unambiguous.
need to check whether this correlates with the degree to which those two predicates applying simultaneously is atypical for the concept. In order to check this, we performed a one-sided correlation analysis between all the results of experiment 1 and those of experiment 2 (figure 2). The result was a positive correlation between mean proportion acceptability of a sentence in a non-interactive interpretation and mean atypicality rating of a predicate pair applying simultaneously ($r = .66, n = 36, p < .001$).

![Figure 2: Relation between mean acceptability (in proportions) of sentences in a 'split' situation and mean atypicality rating of predicate pairs applying simultaneously.](image)

**Discussion**

This paper reports on an experimental investigation into the interpretation of plural sentences with predicate conjunction, and its connection to typicality. We proposed that the extent to which non-interactive interpretations are available directly correlates with the atypicality of a situation where the two predicates apply simultaneously. Experiment 1 revealed a continuum of acceptability values of 36 sentences in a non-interactive, 'split' situation, ranging from 24% to 100% acceptable. Such a continuum is unexpected under the extended SMH by Winter (2001), which assumes that any given sentence is either true or false in a particular situation, depending on what the context allows. Next, experiment 2 showed that different typicality effects exist with respect to complex predicate concepts that are composed of two singular predicates – similar to the effects that were found repeatedly for one-place predicates (e.g. Rosch, 1973). Specifically, we measured the typicality of a situation in which both predicates apply simultaneously, for a given complex concept. The typicality ratings for 36 pairs ranged over the entire 6-point scale. We proposed that typicality relates to acceptability in such a way that the less typical the situation in experiment 2 is judged to be, the more a non-interactive interpretation is available. Based on a correlation analysis, we can conclude that this prediction was borne out. We take these results to be an indication of conceptual structure of predicates playing a crucial role in sentence interpretation, in line with similar results on reciprocal sentences (Struiksma et al., submitted).

An important next step is to further explore typicality effects for complex predicates. In the current paper, we report an experiment that used one particular typicality measurement with one particular dependent measure, namely the typicality of two simultaneous actions, rated on a scale. One can imagine that in fact the typicality of the opposite situation, i.e. two predicates applying to two separate individuals, or perhaps sequentially to one individual, might also affect the interpretation of a plural sentence with those predicates. Also, it will be good to correlate rating measures with different kinds of dependent measures such as categorization speed or error rate to have a more robust result – similar to the investigations into typicality effects for nouns. However, the fact that even one measure can distinguish different types of verb pairs so clearly, is a promising starting point for this enterprise.

Another related issue is the deeper question of how typicality effects come about: What exactly makes a particular instance of a concept typical? A potential candidate factor is that typicality is formed by prior experiences or likelihood of a situation. An anonymous reviewer, however, pointed out an example like (9).

(9) The boys are unicycling and juggling

The reviewer claims that despite the fact that we probably rarely see a person simultaneously unicycling and juggling, we still interpret the conjunction in sentence (9) interactively. Such an example points out that typicality is a complex matter that needs to be studied further. The question of what makes something typical does not affect the results described in this paper per se, but knowing what affects typicality would give them more explanatory power, as pointed out by this reviewer.

Another logical next step would be to investigate other cases in language where typicality affects reasoning with logical operators. So far we have seen that understanding both reciprocal sentences and the sentences with conjunction that were investigated in the current paper, is inseparable from the study of concepts. It is highly likely that these are not the only areas in which this is the case.

**Conclusion**

This paper started from the observation that plural sentences with conjunctive predicates do not always receive the same logical interpretations. Previous work on reciprocal sentences has already taught us that lexical information can influence sentence meaning in systematic ways (e.g. Dalrymple et al., 1998; Kerem et al., 2009; Struiksma et al., submitted). Here we reported on experimental investigation of plural sentences with predicate conjunction, that provided insight into specifically the role of typicality information of predicate concepts. With this result, we add to the line of work that investigates the interface between lexical and compositional semantics, and lead the way towards directions for further research in this area.
Acknowledgments

This work was supported by NWO (The Netherlands Organisation for Scientific Research) VICI grant number 277-80-002, “Between Logic and Common Sense: the Formal Semantics of Words” (2010-15). I am grateful to Hanna de Vries, Marijn Struiksma, Martin Everaert, Choonkyu Lee, Sonia Kasyanenko, Chris Barker, two anonymous reviewers of the “Formal and Experimental Pragmatics” workshop (Tübingen, 2014), and the audiences at the Concept Composition workshop (Utrecht, 2013) and the NYU Semantics Group (New York, 2014) for useful remarks and questions. Special thanks to Yoald Winter for extensive discussions throughout the course of this research. Finally, I am thankful to Nir Kerem for initial ideas that eventually led to the study reported in this paper. The illustration in figure 1 was made by Ruth Noy Shapira.

References


Rosch, E. and C. B. Mervis (1975). Family resemblances: