

# Concept Typicality and the Interpretation of Plural Predicate Conjunction

Eva B. Poortman

**Abstract** This chapter studies the interpretations of plural sentences with conjoined predicates, e.g. *The boys are sitting and cooking* and *The boys are waving and smiling*. Such sentences are sometimes interpreted intersectively, sometimes non-intersectively (or ‘split’), and sometimes both interpretations appear to be allowed. This is surprising, since the logical structure of these sentences is identical, i.e. they differ only with respect to content words (e.g. *sitting, cooking* vs. *waving, smiling*). I propose that the logical interpretation of these sentences is systematically affected by lexical information tied to the complex predicate in the sentences, specifically their so-called typicality effects. With a set of experiments, I show that (a) the acceptability of a sentence in a non-intersective situation 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, I specify at least one of the relevant pragmatic considerations that determine the interpretation of a plural sentence with conjunctive predicates. More generally, these results stress the importance of conceptual structure of predicates in semantic theories of language.

## 1 Introduction

For a long time, it has been a common tradition in logical semantics to draw a clear line between lexical knowledge and compositional operations (Cruse 1986; Jacobson 2014). Logical semanticists generally focus on the latter, studying the way meaning is composed and how syntax affects this process, mostly irrespective of word meanings. Even though lexical semantics is acknowledged to some degree, the connection between the two domains of study has always been relatively weak. Such an approach proved sufficient in the study of the meanings of logical expressions such as *all* or *at most three*. Whether we speak of *all girls*, *all giraffes*

---

E.B. Poortman (✉)

Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, The Netherlands  
e-mail: evapoortman@gmail.com

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3,  
DOI 10.1007/978-3-319-45977-6\_6

or even *all giraffes owned by girls*, does not change the logical contribution of *all* to the meaning of the phrase. One would expect a similar independence of word meaning with logical operators such as *each other* and *and*, e.g. the logical contribution of *and* does not change between *I saw boys and girls* and *I saw giraffes and elephants*. Recently however, studies that looked into these operators revealed that we can no longer make do with such a simple division of labor (e.g. Dalrymple et al. 1998; Kerem et al. 2009; Poortman et al. 2017).<sup>1</sup> These works have shown that contextual information (containing lexical information) and compositional semantics can interact, exposing the need for a more complex story than what is generally assumed. The current chapter deals with one area in which we see ‘logical’ sentence meaning being affected by ‘non-logical’ word meaning, namely plural predicate conjunction, proving that the connection between the two domains is actually fruitful.

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 sitting and each boy is reading, and sentence (2) is true iff each boy is waving and each boy is 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 and Faltz 1985; Partee and Rooth 1983), and combining it with a distributivity operator that shifts a VP into one that holds of a plural individual such as *the boys* iff that VP holds of each atomic part of that individual, i.e. each boy (Link 1983). Importantly, such an analysis assumes that the way we logically reason about these natural language sentences is independent of the lexical elements they contain (such as *sitting*, *standing*, *waving* or *smiling*). 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 chapter, I 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—I will call this a ‘split’ interpretation<sup>2</sup> (Heycock and Zamparelli 2005).

---

<sup>1</sup>In fact, even a seemingly simple term such as *all* has been shown to be context-dependent at times Leslie et al. (2011).

<sup>2</sup>Throughout the chapter, when using the term ‘interpretation’ I simply refer to the situations that support a truthful usage of a sentence.

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 that is tied to the verbs in the sentence, as this appears to be inseparable from a proper analysis of conjunction. I show that there is in fact a continuum of acceptability values (i.e. percentages of “true” judgments) for sentences with conjunctive predicates in non-intersective or ‘split’ situations, and I account for this continuum with a principle that predicts how language users apply predicates to plural subjects based on the typicality structure of the complex predicate.

## 2 Context and Logical Meaning

The insight that context in general, and lexical information in particular, systematically predicts the logical interpretation of a sentence is relatively new. A turning point came with an influential paper on reciprocity by Dalrymple et al. (1998). Dalrymple et al. were the first to start to incorporate a notion of context within logical meaning. They put forward a principle called the Strongest Meaning Hypothesis (SMH), which aims to resolve ambiguity that is caused by contextual information, specifically in the area of reciprocals. Structurally similar reciprocal sentences as in (5) and (6) receive different logical interpretations, despite the fact that they merely differ with respect to the lexical information in the context of the reciprocal expression.

(5) The boys know each other

(6) The boys are following each other

For each occurrence of the reciprocal, the SMH 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. 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 (assuming the boys are not following each other in a circle). Summarizing, we can say that the SMH, unlike what logical semanticists would assume, takes non-logical information to be relevant in determining logical interpretation. It does so without delving too much into what exactly constitutes this non-logical information, other than referring to it as ‘context’ in general. However, without a specific notion of context, empirically supporting such a principle is very difficult.

Nevertheless, Winter (2001) re-uses the main gist of the SMH as a solution to the different interpretations of plural predicate conjunction. Many previous works have described non-intersective interpretations of plural sentences such as example (3) above, as well as sentences with noun phrase conjunction as in (7) given below (e.g. Krifka 1990; Heycock and Zamparelli 2005).

(7) John and Mary wrote an article together

Sentence (7) does not entail that John wrote an article together and Mary wrote an article together, similar to sentence (3) which does not entail that the boys are sitting and those same boys are standing. Krifka (1990) proposes to extend the generally accepted non-intersective conjunction of noun phrases as in (7) to conjunction of predicates as in (3). He proposes that any conjunction  $P_1$  and  $P_2$  holds of an entity  $x$  if  $x$  can be partitioned into two entities  $x_1$  and  $x_2$  such that  $P_1$  holds of  $x_1$  and  $P_2$  holds of  $x_2$ . For (3), this means that whenever the entity ‘the boys’ can be partitioned into two entities, then the predicate *sitting* can hold of one of these entities and the predicate *standing* can hold of the other. Winter (2001) acknowledges that while this is a proper analysis for sentences like (3), it fails to capture the fact that sentences like (1) only allow an intersective interpretation. He claims that on top of Krifka’s descriptive proposal of non-intersective conjunction in addition to intersective conjunction, we also need a principle that determines when which analysis is allowed, thus when the different interpretations actually occur. Winter (2001) proposes that a maximality principle like the SMH is a suitable candidate. First, he assumes that the SMH is not construction-specific to plural sentences with reciprocals. He rephrases it 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). Note that unlike Dalrymple et al.’s SMH, Winter’s extended SMH does not speak of ‘context’ in general, but focuses on a more manageable part of context, namely the lexical information tied to predicates. 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 chapter, I argue that the predictions made by the SMH can be too strong, specifically because its notion of context is still not defined specifically

enough. Consider sentences (8) and (9), which are of a similar nature to sentence (4) above.

(8) The men are lying down and drinking

(9) 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 I will show in the current chapter, non-intersective, 'split' interpretations are readily available to many speakers for sentences like (4), (8) and (9), 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 wave and draw 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 et al. 2009; Poortman et al. 2017). For example, both Kerem et al. (2009) and Poortman et al. (2017) showed experimentally 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*.

### 3 Typicality: Defining Context

These examples, both with reciprocals and conjunction, point to a fundamental issue with the proposal at hand. Since context is not specified in much detail, 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 so-called typicality effects in categorization. The notion of typicality (at least as it is assumed in the current study) simply 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. To illustrate, besides being able to categorize a *sparrow* and an *ostrich* within the *bird* category and a *bat* and a *crocodile* outside of it, people also distinguish between members of a category: e.g. a *sparrow* is judged a more typical bird than an *ostrich*. Since the 1970's, a range of psychological studies has shown for such 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 et al. 1974; Rosch and Mervis 1975). Moreover, I follow Hampton (2007) in assuming that an instance's category membership and its typicality as an instance of that category are two related behavioral measurements, based on one and the same underlying variable. For example, there is a correlation between binary membership measures for *sparrow* (1), *ostrich* (1), *bat* (0) and *crocodile* (0) on the one hand, and their typicality rating on the other hand (*sparrow* > *ostrich* > *bat* > *crocodile*). Hampton assumes a so-called threshold model, according to which there is a threshold somewhere along a typicality function that makes a binary distinction between members (*sparrow*, *ostrich*) and non-members (*bat*, *crocodile*).

For the current purposes, taking into account typicality means extending the aspects of meaning that the SMH is sensitive to from definitional to prototypical, thus fleshing out what constitutes context.<sup>3</sup> Incorporating such 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; Poortman et al. 2017). Firstly, this hypothesis assumes that typicality effects also exist for verb concepts like the binary predicate concept *pinch*, i.e. it assumes that subjects can consistently rank some instances of pinching as more typical than others. The difference with noun concepts like *bird* is thus merely a matter of the *type* of things that are being categorized, namely events instead of objects. Secondly, it predicts that these typicality effects for verb concepts systematically affect the logical interpretation of the reciprocal expression that they combine with. Specifically, the Maximal Typicality Hypothesis (MTH) predicts the core situation for a reciprocal sentence to be the maximal one among those that are most typical for the predicate concept in the sentence (for an elaborate discussion see Poortman et al. 2017).

In this chapter, I extend the same logic to plural sentences with predicate conjunction. I view the MTH as a general principle of meaning composition that systematically governs vagueness in plural sentences. The MTH as such a general mechanism surfaces whenever a graded concept such as reciprocity (in reciprocal sentences) or distributivity (in predicate conjunction sentences)<sup>4</sup> combines with a natural concept such as a verb or an adjective—which each have their own typicality structure. Accordingly, I claim that typicality also affects interpretation in

---

<sup>3</sup>Note that Dalrymple et al.'s notion of context is not very clear, and seems to include the predicate in the scope of the reciprocal as well as things like world knowledge and speaker intentions. All I mean here is that I study one particular aspect of what they refer to as 'context', namely the predicate concepts, and I use typicality as a probe into it.

<sup>4</sup>I assume that reciprocity and distributivity are graded similar to how simple plural sentences are (Winter 2017). Take for example the sentence *The men are sitting*. Such a sentence is more often judged true the more men are actually sitting. Similarly for reciprocal sentences like *The men know each other* (which is more often judged true the more knowing pairs there are) and predicate conjunction sentences like *The men are sitting and cooking* (which is more often judged true the more men are both sitting and cooking).

plural sentences where two predicate concepts are conjoined. Note that my aim here is not to provide a theory of concepts nor to explain how typicality judgments come about (for an explanation of relevant notions of typicality see Hampton (2017)), but merely to investigate the relationship between interpretation and typicality. The proposal works as follows.

We know that predicate conjunctions such as (1) through (4) are classically analyzed intersectively, such that both conjoined predicates apply to each individual in the plural subject simultaneously. The MTH predicts that the degree to which a weaker interpretation is available depends on typicality. Take for example the predicate concept *sitting*, for which we assume typicality effects much like for the concept *bird*: within the instances that are categorized as ‘sitting’ instances (or members of the *sitting* category), I predict that some are consistently judged more typical than others. For example, an event in which a person is sitting straight up in a chair is probably judged as a more typical instance of *sitting* than an event in which a person is leaning so far back that they are almost lying down on the floor. I expect that the different interpretation patterns of sentences (1)–(4) arise due to similar typicality effects with verb concepts. To illustrate, I predict that an event in which a person is sitting while also reading can easily be categorized as a typical instance of the concept *sitting*, and similarly an event in which a person is reading while also sitting can be categorized as a typical instance of the concept *reading*. The fact that both predicates apply simultaneously (i.e. a person is reading while sitting, or sitting while reading) does not affect the typicality of the event for each predicate concept in isolation. For predicate combinations for which this is the case, I predict that plural sentences with combined predicates (e.g. *the boys are sitting and reading* in example (1)) simply behave according to an intersective analysis, i.e. we multiply the number of times the two predicates apply simultaneously based on the number of individuals that the plural refers to, since this does not affect the typicality of the entire situation. By contrast, I predict that events in which a person is sitting while also standing are physically impossible, and therefore not categorized as instances of *sitting*.<sup>5</sup> Similarly, they are not instances of *standing* either. Assuming a threshold model (Hampton 2007), we could say that they fall below the threshold for category membership of *sitting* or *standing* (when the typicality of an event increases for the concept *sitting*, it decreases for the concept *standing*, and vice versa). If we now consider a plural sentence with these predicates combined (e.g. *the boys are sitting and standing* in example (3)), an intersective interpretation is physically impossible (or of close-to-zero typicality for each concept), causing us to weaken the interpretation such that each individual that the plural subject refers to either sits or stands. Crucially, the MTH does not merely make predictions about these two extreme cases, but in fact considers them to be end points on a scale. An example like *the boys are sitting and cooking* (example (4)) clarifies this. Consider

---

<sup>5</sup>Or, if you could imagine some strange situation of sitting and standing simultaneously, then it would at least be a highly atypical instance of *sitting*. This does not affect the nature of the argument.

an event in which a person is sitting while also cooking. Such an event, however odd, can probably be categorized as an instance of the concept *sitting* (i.e. it is above the category membership threshold). And, similarly, an event in which a person is cooking while also sitting can be categorized as an instance of the concept *cooking*. Interestingly however, the fact that both predicates apply at the same time, causes the typicality of the event as an instance of each concept in isolation to decline. An event in which one is sitting while cooking is probably not the most typical instance of *sitting*, and an event in which one is cooking while sitting is definitely not a typical instance of *cooking*.

Crucially, this degree of atypicality that I predict is caused by the simultaneous application of two predicates within one event. Therefore, I will henceforth speak about ‘compatibility’ as a measurement of this atypicality, allowing me to directly compare compatibility between different pairs of predicates. Ideally, if one were to be interested in the full typicality structure of verb concepts like *sitting*, one could construct a standard task (similar to the tasks used for noun concepts like *bird*), namely rating all possible *sitting* events with respect to how typical they are for the concept. In the current chapter however, I restricted my measurements since they are led by a direct research question, based on an observation: I was specifically interested in the different interpretation patterns of sentences like (1)-(4), thus looking for a measurement that allowed me to directly compare “sitting and cooking” versus “sitting and reading” versus “sitting and standing”. One should keep in mind, however, that when I speak of ‘compatibility’ of predicate concepts  $P_1$  and  $P_2$ , I aim to indirectly measure the typicality of an instance of  $P_1$  in an event that has been categorized as an instance of  $P_2$ , and the typicality of an instance of  $P_2$  in an event that has been categorized as an instance of  $P_1$ . The reason I did not test this directly, i.e. presenting subjects with an event in which both predicates apply and measure its typicality as an instance of concept  $P_1$  and of concept  $P_2$  (similar to measuring the typicality of an *ostrich* as an instance of the concept *bird*), was because I was also interested in strictly incompatible pairs of predicates (e.g. *sitting and standing*)—which cannot be depicted within one event. Moreover, I did not use a direct textual test either, i.e. “rate how typical it is for a person to do  $P_1$  in a situation in which she is known to be doing  $P_2$ ” because this seemed to me a harder and more confusing task than a simple compatibility task. Thus instead, I conducted a more indirect, simple textual compatibility test, in which I assess the typicality of an event for concept  $P_1$  and concept  $P_2$  by measuring the predicates’ compatibility.

The measured compatibility is predicted to affect the way sentences are interpreted in which those predicates combine with a plural. Specifically, the less typical the intersective situation is for the two combined predicate concepts in isolation, the more we diverge from an intersective interpretation when those combine with a plural. In more general terms, one could say that when we interpret sentences like (1) through (4) there are two factors at work: (1) maximize the number of predications and (2) retain typicality. These factors are sometimes in conflict, which is when the MTH surfaces. Summarizing, I phrase the proposal as follows:



### MTH for plural predicate conjunction

For any sentence *The X P<sub>1</sub> and P<sub>2</sub>*, in which *X* is a plural and *P<sub>1</sub>* and *P<sub>2</sub>* are singular predicates; and an event *E* in which *P<sub>1</sub>* and *P<sub>2</sub>* apply simultaneously to one individual member of *X*:

The less typical event *E* is for concept *P<sub>1</sub>* and for concept *P<sub>2</sub>*, the more we diverge from an intersective interpretation of the sentence *The X P<sub>1</sub> and P<sub>2</sub>*.

Crucially, this formulation of the MTH assumes that both the notions of typicality and of acceptability can be expressed in terms of a continuum—allowing for more subtle distinctions than the SMH. The experiments that are discussed below measure typicality (via compatibility) and interpretation separately. I predict to find a) that there is a continuum of typicality values for event *E* as an instance of predicate *P<sub>1</sub>* and of predicate *P<sub>2</sub>*, b) that there is a continuum of acceptability values for a plural sentence with those predicates (*The X P<sub>1</sub> and P<sub>2</sub>*) in a given situation and c) that the values on both continuums correlate—indicating that typicality of an event for particular concepts in isolation systematically affects interpretation of sentences containing those concepts. I conducted two behavioral experiments and a correlation analysis to test these predictions.

## 4 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 X are P<sub>1</sub> and P<sub>2</sub>* (where *X* is a plural noun and *P<sub>1</sub>* and *P<sub>2</sub>* are predicates) in a non-intersective, ‘split’ situation. Experiment 2 measured compatibility of predicate concepts *P<sub>1</sub>* and *P<sub>2</sub>* as an indirect typicality test, as argued in the previous section. Materials for the experiments were constructed based on pretests that were conducted in order to include a wide range of compatibility values in the actual experiments.

### 4.1 Pretests: Constructing Materials

The aim of the first pretest was to gather as many Dutch verb combinations as possible, especially atypical ones. I provided 8 participants with 16 sets of two pairs of predicates, *P<sub>1</sub>* and *P<sub>2</sub>* and *P<sub>1</sub>* and *P<sub>3</sub>*: one very natural pair, and one pair that is physically impossible to apply simultaneously, e.g. *sitting and reading* (*P<sub>1</sub>* and *P<sub>2</sub>*) and *sitting and standing* (*P<sub>1</sub>* and *P<sub>3</sub>*). I then asked them to provide as many verbs that they could come up with that combine with *P<sub>1</sub>* (i.e. *sitting* in this case) that led to a possible but atypical, uncommon or strange pair. The pairs that participants constructed, combined with more natural pairs that I came up with, led to a list of 91 verb combinations in total.

In the second pretest, 29 different participants rated all of these 91 pairs for compatibility, in a paper-and-pencil task. For each pair, participants were asked to rate how odd<sup>6</sup> 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’. I mentioned explicitly that 5 thus meant ‘very odd, but physically possible’, in order to distinguish large atypicality from impossibility, or in other words: to distinguish members from non-members by indicating that the category membership threshold for  $P_1$  or  $P_2$  is between 5 and 6. Results of this pretest showed great variability in ratings between verb pairs, with a high level of agreement between the participants (Cronbach’s alpha was 0.88 for the 91 items). The selection of verb pairs that were to be used in Experiments 1 and 2 proceeded as follows. I defined sets of verb pairs on the basis of the different  $P_1$  verbs, e.g. a set consisted of *sitting and reading*, *sitting and standing*, *sitting and knitting*, *sitting and cooking*, etc. Then I selected the 12 sets that showed the greatest range of ratings. Finally, three verb pairs from within each of these 12 sets were selected: 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<sup>7</sup> (atypical pairs like *sitting and cooking*). The 36 verb pairs that constituted the final material, translated from Dutch, are given in Table 1 (the original Dutch material can be found in the Appendix). Creating the three groups (with labels ‘compatible’, ‘incompatible’ and ‘atypical’) was done purely to ensure variability while constructing the materials. I will refer to these three groups when discussing set-up and results of Experiments 1 and 2. Note however that the distinction between the groups is not meaningful in the final correlation analysis of all data points.

#### 4.2 *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 X are  $P_1$  and*

---

<sup>6</sup>Phrasing the question negatively by asking ‘how odd’ subjects would rate a situation was done because (a) directly asking for ‘how compatible’ they would judge two predicates seemed like a too technical and too direct task, and (b) 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.

<sup>7</sup>Additional inclusion criteria included that (a) each verb should be expressed by one word only, (b) ratings for verb pairs should have small variation (whenever there was more than one candidate for selection, the one with the lowest standard deviation for the ratings was selected). Finally, if after considering these criteria there were still two candidate pairs for the atypical group, I decided that (c) atypical verb pairs should have no 6 point ratings (since that meant that at least one participant judged it to be physically impossible for the two verbs to apply simultaneously). This was only a very small criterion, applying to one case.

**Table 1** Overview of predicate pairs, translated from Dutch

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

$P_2$  (where  $X$  is a plural noun and  $P_1$  and  $P_2$  are verbal predicates). The reason for using a ‘split’ situation was that sentences with incompatible pairs cannot be depicted any other way, and I wished to keep all factors in the comparison between pairs equal.

**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 X are  $P_1$  and  $P_2$* )<sup>8</sup> and a drawing depicting four individuals in a non-intersective, ‘split’ interpretation of that sentence: predicate  $P_1$  applied only to persons 1 and 2, predicate  $P_2$  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 Fig. 1.

In each version of the experiment, one third of the test items contained sentences with verb pairs that were considered compatible  $P_1$  and  $P_2$  in the second pretest (e.g. *The men are sitting and reading*), one third contained sentences with verb pairs that were considered incompatible  $P_1$  and  $P_2$  (e.g. *The men are sitting and standing*) and one third contained sentences with pairs that were considered atypical  $P_1$  and  $P_2$  (e.g. *The men are sitting and cooking*). The same drawings were used for sentences with compatible and incompatible pairs with identical  $P_1$  (e.g. *The men*

<sup>8</sup>All the sentences in the experiment were in the simple present tense, which can be used to describe ongoing events as well as states in Dutch. Whereas in English one would use the progressive tense for all sentences in Experiment 1, the distribution of the progressive tense in Dutch is different, such that it could not be used for all sentences in the experiment alike.



**Fig. 1** Example of a test item drawing of Experiment 1. All rights for further use of the illustration arranged with the artist

*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 Fig. 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.

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 P*) or sentences mentioning specific individuals in the picture (*Boys A, B and C are P*). 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 and 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, I 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 et al. 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.

No verb used in the practice session appeared 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. I computed the proportion of true-responses for each of the three types of sentences for each participant. I then performed a repeated measures ANOVA across participants with Compatibility as the within-subjects factor (with 3 levels: compatible, atypical, and incompatible).<sup>9</sup> Post hoc Bonferroni corrected multiple comparisons were performed in order to analyze differences between different Compatibility levels in detail. An ANOVA across items, with Compatibility as the between-item variable (also with 3 levels), gave similar results to the participant analysis. Therefore only the first analysis is reported.

**Results** Table 2 provides an overview of the data. It shows the acceptability of sentences, i.e. the percentage of “true” judgments, for the three levels of Compatibility that were tested for all versions taken together. More detailed results on acceptability per item are in the Appendix. Overall, the truth percentages of the different sentences in the experiment ranged from 24% to 100%. I predicted lowest acceptability for the sentences with compatible pairs and highest acceptability for the sentences with incompatible pairs.

A repeated measures ANOVA revealed that there was a main effect of Compatibility ( $F(1.36, 43.49) = 37.41, p < 0.001$ ). 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 ( $p < 0.001$ ); the acceptability of sentences with compatible predicates differs from the acceptability of sentences with atypical predicates ( $p < 0.001$ ); and the acceptability of sentences with incompatible predicates differs from the acceptability of sentences with atypical predicates ( $p < 0.05$ ). Note again, however, that the main conclusion from

**Table 2** Mean acceptability

Compatibility type	% “true” judgments (st. dev.)
Compatible	54 (31)
Atypical	78 (31)
Incompatible	84 (32)

<sup>9</sup>A repeated measures ANOVA with Version as between-subjects factor was also performed, but showed no effect of Version ( $F(3, 29) = 0.47, p = 0.71$ ) nor an interaction effect of Version \* Compatibility ( $F(6, 58) = 0.82, p = 0.58$ ). I thus collapsed the versions for the analysis.

this experiment is not that there are significant differences between groups, but the fact that there is variation in the data.

### 4.3 *Experiment 2: Compatibility of Predicate Pairs*

This experiment checked compatibility for the 36 predicate concept pairs that were used in sentences of Experiment 1. I aimed to measure the typicality of one particular event as an instance of the concepts  $P_1$  and  $P_2$ , namely one in which both predicates  $P_1$  and  $P_2$  apply simultaneously (event  $E$ ). As already discussed at length earlier on in this paper, I conducted an indirect textual test in which I assess the typicality of event  $E$  for  $P_1$  and  $P_2$  by measuring the predicates' compatibility. This test was identical to the pretest, but carried out by different subjects and now containing fewer items, in a fully controlled experiment.

**Participants** The same 33 students from Utrecht University from Experiment 1 participated 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 and 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, four different orders of the questionnaire were constructed: 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 odd<sup>10</sup> they would consider it if both verbs applied to the given person at the same time. Oddness was rated on a

---

<sup>10</sup>As mentioned in footnote 6, phrasing the question negatively by asking 'how odd' subjects would rate a situation was done because directly asking for 'how compatible' or 'how typical' they would judge a situation turned out to be unsuitable.

6-point scale, where 1 meant ‘not odd at all’ and 6 meant ‘physically impossible’, or in other words: to distinguish members from non-members by indicating that the category membership threshold for  $P_1$  or  $P_2$  is between 5 and 6. It was 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 oddness judgment. This way the incompatibility rating for each verb pair was computed. I 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 incompatibility rating for the three levels of Compatibility that were tested, for all versions taken together. More detailed results on incompatibility rating per item are in the Appendix. Overall, mean ratings per verb pair ranged from 1.03 to 5.94, and there was a very high correlation between these ratings and the ratings for these items in the pretest ( $r = 0.98, p < 0.001$ ).

A repeated measures ANOVA revealed that there was again a main effect of Compatibility ( $F(1.95, 62.45) = 1187.02, p < 0.001$ ). This means that the mean incompatibility ratings for the three Compatibility levels are not equal. Pairwise comparisons show that all three levels differ significantly from each other: the incompatibility of supposed compatible pairs differs from the incompatibility of supposed incompatible pairs ( $p < 0.001$ ); similarly for the incompatibility of compatible vs. atypical pairs ( $p < 0.001$ ); and the incompatibility of incompatible vs. atypical pairs ( $p < 0.001$ ). This means that the three groups that were selected based on the pretest were confirmed in Experiment 2 (with different subjects and a subset of the stimuli).

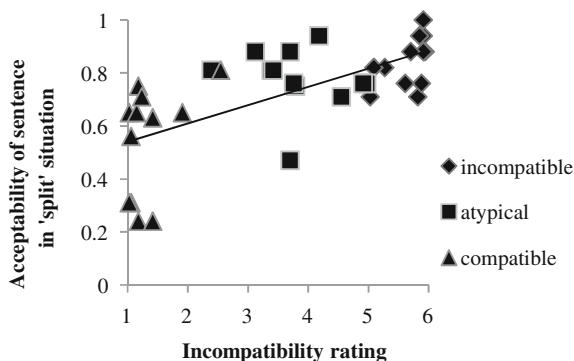
#### 4.4 Correlation Between Interpretation and Compatibility

The crucial test for the proposal is the relationship between interpretation and compatibility. In order to account for the degree to which non-intersective interpretations of sentences *The X are  $P_1$  and  $P_2$*  are available given two particular conjoined predicates  $P_1$  and  $P_2$ , we need to check whether this correlates with the degree to which  $P_1$  and  $P_2$  are incompatible (as an indirect measurement of the atypicality of  $P_1$  and  $P_2$  applying simultaneously (event  $E$ ) for each concept in

**Table 3** Mean incompatibility ratings

Compatibility type	Mean incompatibility (st. dv.)
Compatible	1.35 (0.29)
Atypical	3.82 (0.58)
Incompatible	5.66 (0.38)

**Fig. 2** Relation between mean acceptability (in proportions) of sentences in a ‘split’ situation and mean incompatibility rating of predicate pairs.



isolation). In order to check this, I performed a correlation analysis between all the results of Experiment 1 and those of Experiment 2 (Fig. 2). The result was a positive correlation between mean proportion acceptability of a sentence in a non-intersective interpretation and mean incompatibility rating of a predicate pair ( $r = 0.66$ ,  $n = 36$ ,  $p < 0.001$ ).

## 5 Discussion

This paper reports on an experimental investigation into the interpretation of plural sentences with predicate conjunction, and its connection to typicality. I proposed that the extent to which non-intersective interpretations are available directly correlates with the atypicality of an event in which the two predicates apply simultaneously. Experiment 1 revealed a continuum of acceptability values of 36 sentences in a non-intersective, ‘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 differences in compatibility exist between different predicate pairs. The compatibility ratings for 36 pairs ranged over the entire 6-point scale. I assumed that the compatibility measurement is an indirect measurement of typicality, namely of the typicality of event  $E$  (in which two predicates apply simultaneously) for each predicate concept in isolation, and hence that this effect is similar to the effects that were found repeatedly for one-place predicates (e.g. Rosch 1973). I proposed to extend the Maximal Typicality Hypothesis (Kerem et al. 2009; Poortman et al. 2017) by formulating it for predicate conjunction in such a way that typicality relates to acceptability so that the less compatible the two predicates in Experiment 2 are judged to be (i.e. the less typical event  $E$  is), the more a non-intersective interpretation is available. Based on a correlation analysis, I conclude that this prediction was borne out. Note that this correlation does not hinge on my assumption that compatibility is an indirect way of measuring typicality. I merely



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 (Poortman et al. 2017).

### 5.1 Reference Shift of the Plural Subject?

The particular interpretation that was in the focus of the current study was the so-called ‘split’ interpretation in which  $P_1$  always applies to two of the individuals in the picture and  $P_2$  always applies to the two remaining individuals. I have claimed that this interpretation is sometimes available for predicate conjunction sentences, namely to the degree that a situation in which the conjoined predicates apply simultaneously is atypical. One might argue instead that the acceptability of these sentences given a split interpretation has nothing to do with typicality. As an alternative, one might reason that we accept a sentence like (3) (repeated below as (10)) because its deep structure is the sentential conjunction in (11), which contains two definite plurals that hence allow the possibility of referring to two different groups of boys. In other words, the reasoning would be that we accept (10) in a split situation because we are able to very quickly shift the reference of the plural noun *the boys* from one set of boys to another set of boys. My experiments would then in fact deal with reference resolution instead of with matters of typicality.

(10) The boys are sitting and standing

(11) The boys<sub>i</sub> are sitting and the boys<sub>j</sub> are standing

If we indeed actually interpret the predicate conjunction in (10) as sentential conjunction (as is made explicit in (11)), then I would expect to see no differences between different test sentences. If reference shift would explain why sentence (10) is accepted in a split situation, then we should be able to use this strategy across the board for all types of predicate conjunction that were tested—whether they are typical, atypical or incompatible. This is clearly not the case, and the question remains what explains the range of acceptability values.

In fact, a pilot study<sup>11</sup> has revealed that when sentential conjunction is explicit in the surface form of the sentence (i.e. when subjects are given sentences like (11)), we see that indeed it is possible to shift the referent for different types of predicate

---

<sup>11</sup>The study was conducted with 9 participants who were students at Utrecht University (6 female, age  $M = 23$ ), and checked the acceptability of 12 plural sentences with sentential conjunction in a ‘split’ situation. Each sentence was of the form *The x are  $P_1$  and the x are  $P_2$*  (where  $x$  is a plural noun (used twice) and  $P_1$  and  $P_2$  are verbal predicates). Half of the  $P_1$  and  $P_2$  pairs were compatible predicates while the other half were incompatible predicates (based on pretests from the study reported in this paper). One participant accepted none of the sentences, the remaining eight participants accepted all or all but one.

conjunction. Subjects accepted sentence (11) (with incompatible predicates) given a split situation, but they also accepted sentences like (12) and (13) given a split situation—even though we have seen that sentences containing such compatible predicates behave differently when presented as mere predicate conjunctions (i.e. they are generally judged false 50% of the time). A single subject did not accept sentence (11) in a split situation, but they also refused to accept sentences like (12) and (13) in a split situation.

(12) The boys are sitting and the boys are reading

(13) The boys are waving and the boys are smiling

What this suggests is that reference shift is independent from the obtained results in the current study. Sentences with sentential conjunction do not show the same correlation with typicality as sentences with predicate conjunction do. I believe reference shift of the plural subject in sentences like (11)–(13) is purely motivated by trying to make a sentence true. This explains why most participants always accepted such sentences in a split situation (i.e. used reference shift to make the sentence true). The one subject that did not use reference shift, was consistent in not using it across different types of predicate conjunction (compatible and incompatible).

Another argument against the reference shift explanation is the finding (based on a small pilot<sup>12</sup>) that sentences with proper name conjunctions instead of definite plurals are accepted in a split situation significantly more often when the conjoined predicates are incompatible (as in (14)) compared to when they are compatible (as in (15)).

(14) John, Bill, Sue and Jane are sitting and standing

(15) John, Bill, Sue and Jane are sitting and reading

For such sentences, reference shift of the subject *John, Bill, Sue and Jane* is obviously not possible, and still a sensitivity to the predicate concepts in the sentence is observed, which is along the same lines as the results presented in the current chapter.

Summarizing, I conclude that it is unlikely that the presented results are due to reference shift of the plural subject. Shifting the reference of the subjects from one referent to another does not explain the systematic variability in acceptability, nor that a similar pattern arises for sentences with proper name conjunction as in (14) and (15).

---

<sup>12</sup>This study was conducted with 22 participants who were students at Utrecht University (18 female, age  $M = 19$ ). It checked the acceptability of 8 plural sentences with predicate conjunction in a ‘split’ situation. Each sentence was of the form *A, B, C and D are  $P_1$  and  $P_2$*  (where *A, B, C* and *D* are names and  $P_1$  and  $P_2$  are verbal predicates). Half of the  $P_1$  and  $P_2$  pairs were compatible predicates while the other half were incompatible predicates (based on introspection). Sentences with compatible pairs were accepted in a split situation 10% of the time, sentences with incompatible predicates were accepted in a split situation 40% of the time.

## 5.2 *Other Measures of Typicality*

Despite the fact that we can safely rule out reference shift as an alternative explanation of the results, obviously there are many other factors that are worth further exploration. The correlation that was found in this study was high ( $r = 0.66$ ,  $n = 36$ ,  $p < 0.001$ ), though obviously not perfect. This means that there must be more factors that affect interpretation besides the one tested here. An important next step is to delve deeper into typicality effects for complex predicates. In the current chapter, I report an experiment that indirectly measured 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 event, 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. Moreover, as pointed out by a reviewer, perhaps not only the verb concepts but also the head noun of the sentences play a role. It might be that the compatibility of two predicates is quite different in the context of humans than it is for example in the context of dogs: people can run and scratch their heads simultaneously, but dogs cannot. In order to fully understand the factors that influence sentence interpretation, an intricate combination of typicality measures is necessary.

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 example (16).

(16) 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 probably interpret the conjunction in sentence (16) intersectively (though of course a full sample of participants would need to be consulted to be sure). Such an example points out that typicality is not simply a matter of frequency, but a far more complex notion that needs to be studied further. The question of what makes something typical does not affect the results described in this chapter per se, but knowing what affects typicality would give them more explanatory power, as pointed out by this reviewer.

### 5.3 Further Areas

Another logical step would be to investigate other cases in language where typicality affects reasoning. So far we have seen that understanding both reciprocal sentences and the sentences with conjunction that were investigated in the current chapter, is inseparable from the study of concepts. Another area where we see typicality affecting interpretation, is the area of adjective-noun constructions such as *red hair* (Lee 2017). For such a construction, the typicality structure of *hair* appears to interact with the way we interpret the adjective *red*. Even though the concept *red* in isolation might have as its most typical instance a focal red, orange-like hues are generally more typical for the concept *hair*. When the two combine, these typicality preferences interact (for more on these effects see the work by Lee (2017) and Winter (2017)). This interaction is intuitively of a similar nature to the one between a verb concept like *pinch* and the reciprocal expression *each other*, as well as the one between verb concepts like *sitting* and *cooking* and the logical expression *and*. It is highly likely that these are not the only areas in which this is the case, thus it is worthwhile for further research to investigate whether a principle like the MTH can function as a general principle of language use.

## 6 Conclusion

This chapter 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; Poortman et al. 2017). Here I 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, I 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.

**Acknowledgements** 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” (2010–16). This chapter formed the basis for Chap. 3 of my doctoral dissertation entitled ‘Concepts and Plural Predication: The Effects of Conceptual Knowledge on the Interpretation of Reciprocal and Conjunctive Plural Constructions’, published by LOT (Utrecht, The Netherlands). I am grateful to Hanna de Vries, Marijn Struiksma, James Hampton, Martin Everaert, Choonyu Lee, Sonia Kasyanenko, Chris Barker, two anonymous reviewers of the ESSLII workshop “Formal and Experimental Pragmatics”, two anonymous reviewers of this volume, and the audiences at the Concept Composition workshop (Utrecht 2013), the NYU

Semantics Group (New York 2014) and the Formal and Experimental Pragmatics workshop (Tübingen 2014) for useful remarks and questions. Special thanks to Yoad 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 Fig. 1 was made by Ruth Noy Shapira.

## Appendix

This appendix contains the original Dutch stimulus verb pairs  $P_1$  and  $P_2$  that were used in Experiments 1 and 2, and their English translation. In Experiment 1, each verb pair was contained in a sentence of the form *The X are  $P_1$  and  $P_2$*  (where  $X$  is a plural noun). The two rightmost columns in the appendix give the mean results of each experiment for each stimulus verb pair.

Group	Verb pair $P_1$ and $P_2$	English translation	Experiment 1 Interpretation (% true)	Experiment 2 incompatibility (rating 1–6)
incomp	Zitten en staan	Sitting and standing	0.94	5.91
incomp	Zwaaien en klappen	Waving and clapping	0.82	5.27
incomp	Lopen en zwemmen	Walking and swimming	0.88	5.94
incomp	Kruipen en springen	Crawling and jumping	0.71	5.82
incomp	Staan en hurken	Standing and squatting	0.94	5.85
incomp	Lezen en slapen	Reading and sleeping	0.76	5.88
incomp	Liggen en rennen	Lying down and running	0.88	5.91
incomp	Tekenen en typen	Drawing and typing	0.82	5.09
incomp	Zwemmen en kruipen	Swimming and crawling	1	5.91
incomp	Smsen en breien	Texting and knitting	0.71	5.03
incomp	Breien en klappen	Knitting and clapping	0.88	5.7
incomp	Slapen en bellen	Sleeping and telephoning	0.76	5.61
	<b>MEAN</b>		<b>0.84</b>	<b>5.66</b>
atyp	Zitten en koken	Sitting and cooking	0.88	3.12

(continued)

(continued)

Group	Verb pair $P_1$ and $P_2$	English translation	Experiment 1 Interpretation (% true)	Experiment 2 incompatibility (rating 1–6)
atyp	Zwaaien en tekenen	Waving and drawing	0.75	3.79
atyp	Lopen en schrijven	Walking and writing	0.81	3.39
atyp	Kruipen en lezen	Crawling and reading	0.71	4.55
atyp	Staan en dutten	Standing and falling asleep	0.47	3.7
atyp	Lezen en tekenen	Reading and drawing	0.94	4.18
atyp	Liggen en drinken	Lying down and drinking	0.81	3.42
atyp	Tekenen en lopen	Drawing and walking	0.76	3.76
atyp	Zwemmen en lezen	Swimming and reading	0.76	4.97
atyp	Smsen en zwaaien	Texting and waving	0.81	2.39
atyp	Breien en lopen	Knitting and walking	0.88	3.7
atyp	Slapen en staan	Sleeping and standing	0.76	4.91
	<b>MEAN</b>		<b>0.78</b>	<b>3.82</b>
comp	Zitten en lezen	Sitting and reading	0.65	1.03
comp	Zwaaien en lachen	Waving and smiling	0.31	1.06
comp	Lopen en zingen	Walking and singing	0.24	1.18
comp	Kruipen en schreeuwen	Crawling and screaming	0.81	2.55
comp	Staan en lezen	Standing and reading	0.65	1.91
comp	Lezen en glimlachen	Reading and smiling	0.31	1.03
comp	Liggen en rekken	Lying down and stretching	0.65	1.15
comp	Tekenen en gapen	Drawing and yawning	0.75	1.18
comp	Zwemmen en lachen	Swimming and smiling	0.24	1.42
comp	Smsen en fronsen	Texting and frowning	0.56	1.06

(continued)

(continued)

Group	Verb pair $P_1$ and $P_2$	English translation	Experiment 1 Interpretation (% true)	Experiment 2 incompatibility (rating 1–6)
comp	Breien en zingen	Knitting and singing	0.71	1.24
comp	Slapen en kwijlen	Sleeping and drooling	0.63	1.42
	<b>MEAN</b>		<b>0.54</b>	<b>1.35</b>

## References

- van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, 38(4), 584–589.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge, U.K.: Cambridge University Press.
- Dalrymple, M., Kanazawa, M., Kim, Y., Mchombo, S., & Peters, S. (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy*, 21, 159–210.
- Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science*, 31(3), 355–384.
- Hampton, J. A. (2017). Compositionality and concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 95–122). London: Springer.
- Heycock, C., & Zamparelli, R. (2005). Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural Language Semantics*, 13, 201–270.
- Jacobson, P. (2014). *Compositional semantics*. Oxford, U.K.: Oxford University Press.
- Keenan, E., & Faltz, L. (1985). *Boolean semantics for natural language*. Dordrecht: D. Reidel.
- Kerem, N., Friedmann, N., & Winter, Y. (2009). Typicality effects and the logic of reciprocity. In *Proceedings of Semantics and Linguistic Theory, SALT19* (pp. 257–274).
- Krifka, M. (1990). Boolean and non-boolean ‘and’. In L. Kálmán & L. Pólós (Eds.), *Papers from the second symposium of logic and language*. Budapest: Akadémiai Kiadó.
- Lee, C. (2017). Typicality knowledge and the interpretation of adjectives. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 123–138). London: Springer.
- Leslie, S., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, 65, 15–31.
- Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bauerle et al. (Eds.), *Meaning, use and interpretation of language*. Berlin: De Gruyter.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314–324.
- Partee, B., & Rooth, M. (1983). Generalized conjunction and type ambiguity. In R. Bauerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use and interpretation of language*. Berlin: De Gruyter.
- Philip, W. (2000). Adult and child understanding of simple reciprocal sentences. *Language*, 76(1), 1–27.
- Poortman, E. B., Struiksma, M. E., Kerem, N., Friedmann, N., & Winter, Y. (2017). Reciprocal expressions and the Maximal Typicality Hypothesis. To appear in *Glossa*.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.

- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 81(3), 214–241.
- Winter, Y. (2001). Plural predication and the strongest meaning hypothesis. *Journal of Semantics*, 18, 333–365.
- Winter, Y. (2017). Critical Typicality: Truth judgements and compositionality with plurals and other gradable concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp 163–190). London: Springer.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

