

# Modification and Default Inheritance



Corina Ströbner, Annika Schuster, and Gerhard Schurz

**Abstract** Modification usually decreases the judged likelihood of typicality statements. People judge “Old coyotes howl” as less likely than just “Coyotes howl”. This paper addresses this so-called modification effect. In order to analyse the effect, we propose an extended modification model based on the selective modification model by Smith et al. (1988) and Barsalou’s (1992) frames. In this model we introduce cross-attributional constraints that explain how a change in one dimension leads to an alteration of another attribute, especially if the modifier is not typical. Finally, we discuss data from Connolly et al. (2007) and present new experimental evidence from an explorative study.

**Keywords** Modifier effect · Constraints · Frames · Prototype theory · Compositionality

## 1 Prototype Compositionality and Modification

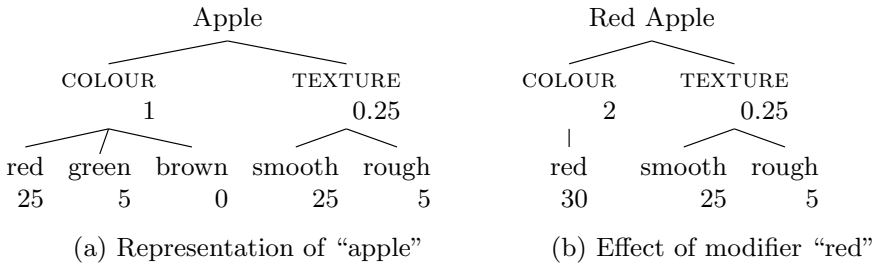
Originating in the work of Eleanor Rosch and her co-authors (Rosch and Mervis 1975; Rosch et al. 1976; Rosch 1978), the prototype theory of concepts influenced the way psychologists, linguists and philosophers understand concepts enormously. In its most popular version, prototype theory claims that concepts are associated with internal typicality orderings. This thesis is well-confirmed and widely accepted. It is also well-known that human agents are capable of composing concepts. But how can prototype theory contribute to the understanding of this creative process? Typicality doesn’t combine in a straightforward way. A typical pet fish is neither a typical fish nor a typical pet (cf. Fodor and Lepore 1996). Elaborated models of prototype composition have been developed since the 1980s. Hampton (1987) discusses how

---

C. Ströbner (✉)  
Ruhr University Bochum, Bochum, Germany  
e-mail: [Corina.Stroessner@ruhr-uni-bochum.de](mailto:Corina.Stroessner@ruhr-uni-bochum.de)

A. Schuster · G. Schurz  
Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

© The Author(s) 2021  
S. Löbner et al. (eds.), *Concepts, Frames and Cascades in Semantics, Cognition and Ontology*, Language, Cognition, and Mind 7,  
[https://doi.org/10.1007/978-3-030-50200-3\\_14](https://doi.org/10.1007/978-3-030-50200-3_14)



**Fig. 1** Modification in SMM, following (Smith et al. 1988, pp. 490, 494)

the typicality ratings of noun constructions like “sports that are also games” are determined by the importance of the properties for their components. The selective modification model proposed by Smith et al. (1988), on the other hand, concerns modifications that are realized in adjective-noun combinations. These are at the focus of this paper.

The selective modification model, henceforth SMM, starts with a representation of prototype concepts as attribute-value structures, an importance measure for attributes, called diagnosticity,<sup>1</sup> and a voting for the values, called salience (cf. Smith et al. 1988, p. 489).

Modification is understood as a strictly selective process in the SMM, the effect of which is limited to one attribute. The modifier selects the attribute the adjective addresses, shifts all votes to its value, and increases the importance of this particular attribute (cf. Smith et al. 1988, p. 492). Figure 1 shows how the modifier “red” for “apple” operates on the colour attribute: all votes go to “red” and the importance of colour is increased. The SMM is a very simple but still effective approach to prototype compositionality. However, it also has several limitations. The most important one is the strictness of its selectivity. This strong assumption prevents modification from altering anything but one attribute. Thus, SMM predicts that all non-modified properties are inherited. Smith et al. (1988, p. 497) are aware of possible correlations between attributes, but they defer necessary adjustments to a subsequent cognitive process.

Connolly et al. (2007) presented experimental evidence that is not compatible with the predictions of SMM: subjects rate unmodified statements like “Ravens are black” more likely than modified ones like “Feathered ravens are black”. The judged likelihood is even lower if the modifiers are not typical, e.g., for “Jungle ravens are black”, and further decreases if two modifiers are used, as in “Young jungle ravens are black”. On a scale from 1 (very unlikely) to 10 (very likely) the mean rating was 8.36 for unmodified sentences (A), 7.71 for typical modifications (B), 6.91 for non-typical ones (C) and only 6.48 for double modifications (D) (cf. Connolly et al. 2007, p. 11f). Jönsson and Hampton (2012) and Hampton et al. (2011) confirmed

<sup>1</sup>The term “diagnosticity” is also often used to indicate the specificity of a property. Thus we prefer the expression “attribute importance”.

these findings in further experiments. Gagné and Spalding (2011, 2014) observed a similar effect for meaningless pronounceable modifiers. While the existence of the effect is uncontroversial, there is a lively debate on its interpretation.

Connolly et al. (2007) claim that their experiment proves that people don't use a default to prototype strategy. Typical values are not inherited by subcategories but rather inferred in post-compositional step, which is largely lead by personal knowledge (cf. Connolly et al. 2007, p. 15). Jönsson and Hampton (2012, p. 109), on the contrary, argue that typical properties are inherited. Only in a second step, subjects decrease their certainty about the typical properties, mostly because of background knowledge or for pragmatic reasons. Gagné and Spalding (2014) take a third position. According to them, typical properties are not inherited but inferred from the meta-knowledge that subcategories resemble the category to some degree but are still distinct (cf. Gagné and Spalding 2014, p. 1291).

The authors agree in taking their results to be incompatible with the SMM. To begin with, it is hard to explain why modifications influence typical values of attributes that are not addressed by the modification. On top of that, the remarkable difference between typical modifiers and other modifiers is an unexplainable mystery to the SMM (cf. Jönsson and Hampton 2012, p. 111). However, the SMM also explains some results. The rating of a modified sentence is highly correlated with the typicality rating of its unmodified counterpart (cf. Jönsson and Hampton 2012, p. 98). The contribution of the head noun to the modification occurs even if the modifier is not a meaningful word, and thus certainly not learned from experience (cf. Gagné and Spalding 2014, pp. 1287–1288). In sum, experimental evidence has revealed three stable effects for a head noun S, its prototypical property P and the modifier M:

1. "S is P" is usually rated as more likely than "M S is P"
2. There is a positive correlation between the rated likelihood of "S is P" and "M S is P".
3. The loss in rated likelihood from "S is P" to "M S is P" is smaller if M is typical for S.

The modification effect doesn't depend on how central the typical property P is. Hampton et al. (2011) even produced the effect for properties that are analytically true, like being a bird for raven.

While a reference to post-compositional adjustments can save the basic idea of SMM, it also reduces its empirical content and strength. This is why we propose to enrich the SMM by making use of frames (Barsalou 1992), i.e. recursive attribute-value structures that allow the specification of constraints between values of different attributes. This is carried out in the next section. In the third section, we show that an application of our model to data from Connolly et al. (2007) shows a stronger decrease in likelihood in the presence of constraints. We finally present new experimental evidence we gathered in an exploratory study.

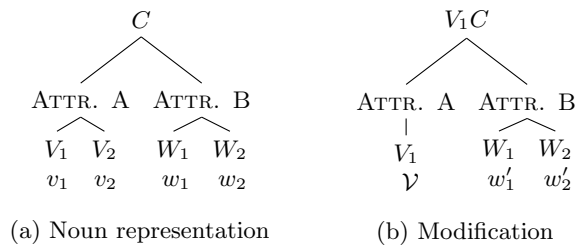
## 2 An Extended Modification Model

We understand modification as an asymmetrical composition that is usually realized in adjective-noun compounds. Depending on the way the modifier interacts with the head noun, normal modifications can be distinguished from deviant forms of modification. In a normal modification, the modifier picks a value for a attribute in the noun frame. Deviant and privative modifications like “stone lion”, on the other hand, are grammatically like normal modifications but interfere with the noun in a more drastic way. They often lead to coercion, metaphorical use or to a high demand for context. When we are confronted with such compounds we have to reconsider our normal interpretation of the head noun. Understanding of deviant modifications confronts us with its own obstacles, the reasons for which do not lie in prototype theory. Our approach is therefore focused on the understanding of normal modifications.

For our illustration of modification we refer to the frame model of Barsalou (1992), who claims that conceptual content is best represented in terms of attribute-value structures. Cross-attributional dependencies are illustrated as constraints, i.e. as relations between values. A constraint can for example state that a green colour of an apple indicates sour taste. Barsalou’s frames comprise the attribute-value structure of SMM and, additionally, they allow the representation of dependencies between values of different attributes by means of constraints.

Our enriched model of modification states, like SMM, that a modifier specifies a value of a noun’s attribute and shifts all votes to this value. SMM also claims an importance boost of the according attribute. Although we readily accept this thesis, it will not matter in the following discussion. Our focus is on the changed likelihood of values, i.e. the shift of votes. Thus, we ignore importance measures in this paper. Our essential extension of the SMM is the *constraint thesis*, which contradicts the strict selectivity of SMM: by modification, the selected value collects all votes *and activates constraints to other values*. The constraint thesis will be formalised in the next section. The discussion is based on the minimal model, shown in Fig. 2. We consider a concept  $C$  with two attributes, A and B. The values of A are  $V_1$  and  $V_2$  with the respective votes  $v_1$  and  $v_2$ . The attribute B has the values  $W_1$  with  $w_1$  votes and  $W_2$  with  $w_2$  votes (Fig. 2a).  $\mathcal{V} = v_1 + v_2 = w_1 + w_2$  is the number of total votes.

Fig. 2 Basic model



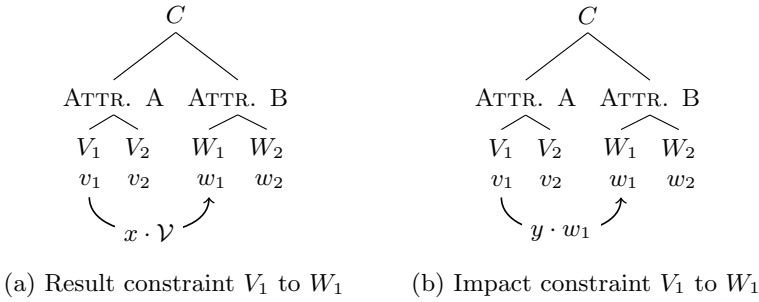


Fig. 3 Constraints

$V_1$  is the value of the modifier, as shown in Fig. 2b, with the new votes  $v'_1 = \mathcal{V}$  for  $V_1$ ,  $v'_2 = 0$  for  $V_2$  and the new votes  $w'_1$  for  $W_1$  as well as  $w'_2$  for  $W_2$ .

**Typicality of values and modifiers**

The typicality of a modifier has a well-documented influence in all experiments. The existing literature, starting from Connolly et al. (2007), distinguishes between modifiers that are typical and other modifiers.<sup>2</sup> We will refine the notion of typicality. Drawing on the distributions of votes on an attribute, we distinguish *typical* values with a very high proportion of votes from *atypical* values with a very low proportion of votes. Values with a medium number of votes are called *neutral*.

**Bringing constraints to SMM**

Since the SMM is based on quantitative specifications, especially votes for values, it is necessary to quantify constraints as well. There are different ways to achieve this.

One possibility to quantify constraints is to specify the proportion of votes that will be given to the target value if the constraining value comes to be known. An example of such a result constraint is given in Fig. 3a:  $x$  with  $0 \leq x \leq 1$  is the proportion of votes the result constraint from  $V_1$  gives to  $W_1$ . The value  $x$  can be interpreted as the conditional probability of  $W_1$  given  $V_1$ . This allows us to tie on to results in the field of probability theory.<sup>3</sup> If no further constraint is involved, the other votes on B ( $w_2, w_3, \dots, w_n$ ) need to be adjusted in a way that reflects their initial proportion, namely as  $w'_i = \frac{w_i}{\mathcal{V}} \cdot (\mathcal{V} - w'_1)$ . The strength or impact of the constraint is apparent by the difference between the initial votes and the new votes.

<sup>2</sup>They refer to typical modifiers as those properties that were collected in feature lists by Cree and McRae (2003).

<sup>3</sup>An elaborated way to model dependencies probabilistically is the theory of Bayesian nets introduced by Pearl (1998). For our basic model, which is based on correlations, Bayesian nets are overly powerful.

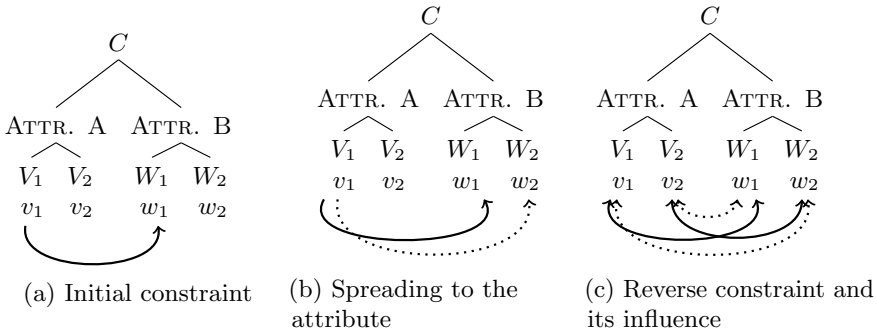


Fig. 4 Influence of constraints

An alternative approach are impact constraints that specify the alteration of the votes by the particular constraint. In the impact representation of a constraint from  $V_1$  to  $W_1$ , we give a factor  $y$  such that  $0 \leq y \leq \frac{1}{w_1/v}$ , which is multiplied with  $w_1$  as illustrated in Fig. 3b. The new votes for  $w_2, w_3, \dots, w_n$  after activating the constraint from  $V_1$  to  $W_1$  are calculated as  $w'_i = w_i \cdot \frac{v-w'_i}{v-w_1}$ . The direction of the constraint is now apparent in the constraint itself: For positive constraints we have  $y > 1$ , while for negative constraints  $y < 1$ . Neutral constraints with  $y = 1$  can be used to represent known irrelevance.

The influence of constraints spreads. Any active constraint from  $V_1$  to  $W_1$  has an influence on  $W_1$ 's alternatives. If  $V_1$  increases the likelihood of  $W_1$ , then it decreases the likelihood of its alternatives, e.g.,  $W_2$ , and the other way around. Furthermore, the constraint from  $V_1$  to  $W_1$  leads to a constraint from  $W_1$  to  $V_1$ . It can be calculated by Bayes' theorem as  $P(V_1|W_1) = \frac{x \cdot (v_1/v)}{w_1/v}$ . Thus,  $W_1$  increases the likelihood of  $V_1$  and decreases the likelihood of  $V_2$  if and only if  $V_1$  increases the likelihood of  $W_1$ . This is illustrated in Fig. 4b, where solid arrows indicate increased likelihoods (positive constraints) and dotted ones indicate decreased likelihoods (negative constraints).

### Constraining constraints

Constraints are restricted. For example, a typical value cannot severely increase the likelihood of an atypical one. In order to determine the impact of a constraint, we introduce the factor  $f$  that is needed to shift all votes to the constraining value  $V_1$ , i.e. to make it maximally probable:  $f = \frac{1}{P(V_1)}$ . To approach the possible influence on  $W_1$ , we rely on  $P(W_1)' = f \cdot P(W_1 \wedge V_1) + 0 \cdot P(W_1 \wedge \neg V_1)$ . For a positive constraint, we stipulate that  $P(W_1 \wedge V_1)$  is as high as possible. Thus,  $f = \frac{1}{P(V_1)}$  is also the maximal positive impact a constraint from  $V_1$  can have, of course still with the limit that  $P(V_1) \cdot f \leq 1$ . For calculating the maximal negative constraint we assume  $W_1 \wedge \neg V_1$  to be as likely as possible.  $P(W_1 \wedge \neg V_1)$  cannot be larger than  $P(\neg V_1) = 1 - P(V_1)$ , i.e.  $P(W_1 \wedge V_1) \geq P(W_1) - P(V_1)$ .

**Table 1** Possible constraints and their results

Typicality of values	$v_1/\mathcal{V}$	$w_1/\mathcal{V}$	$P(V_1)$	$P(W_1)$	$P(W_1 V_1)$		Maximal votes change	
					Max	Min	Gain	Loss
Atypical to atypical	3/30	3/30	$\frac{1}{10}$	$\frac{1}{10}$	1	0	+27	-3
Atypical to neutral	3/30	15/30	$\frac{1}{10}$	$\frac{1}{2}$	1	0	+15	-15
Atypical to typical	3/30	27/30	$\frac{1}{10}$	$\frac{9}{10}$	1	0	+3	-27
Neutral to atypical	15/30	3/30	$\frac{1}{2}$	$\frac{1}{10}$	$\frac{1}{5}$	0	+3	-3
Neutral to neutral	15/30	15/30	$\frac{1}{2}$	$\frac{1}{2}$	1	0	+15	-15
Neutral to typical	15/30	27/30	$\frac{1}{2}$	$\frac{9}{10}$	1	$\frac{4}{5}$	+3	-3
Typical to atypical	27/30	3/30	$\frac{9}{10}$	$\frac{1}{10}$	$\frac{1}{9}$	0	+0.33	-3
Typical to neutral	27/30	15/30	$\frac{9}{10}$	$\frac{1}{2}$	$\frac{5}{9}$	$\frac{4}{9}$	+1.66	-1.66
Typical to typical	27/30	27/30	$\frac{9}{10}$	$\frac{9}{10}$	1	$\frac{8}{9}$	+3	-0.33

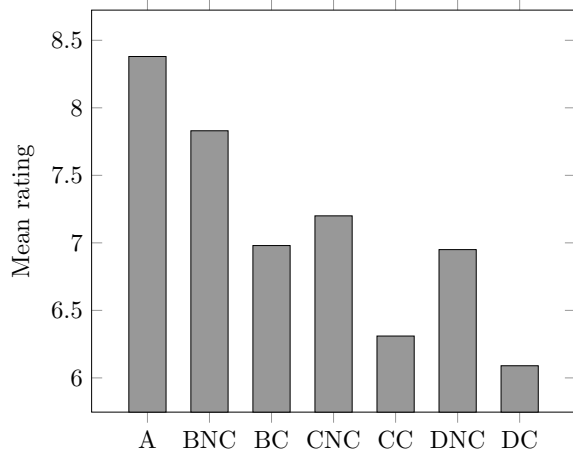
Table 1 shows how the rules restrict the effect of constraints. The initial votes on the modifying value  $V_1$  play a crucial role. If  $V_1$  is rather atypical, i.e. in the first three combinations, then the constraint can change the new distribution of votes severely. A typical modifier  $V_1$ , on the other hand, has only a limited potential to alter the initial distribution of votes.

Besides the formal considerations, there are conceptual restrictions. Prototype concepts represent property clusters (Rosch and Mervis 1975; Schurz 2012). Within the supercategory, typical values of a prototype concepts are positively correlated with each other. This correlation is not always inherited by the subcategories: Within the class of vertebrates, a beak is a good predictor of flying-ability but not in the category of birds. However, the positive correlation often remains valid, if functionality is involved. The beat of the heart is causally related with almost all vital properties of an organism and thus also statistically correlated with them. The typical shape of a tool is adjusted to its typical purposes. Positive associations between typical values in a category are frequent. For the formal reasons explicated above, these constraints should not be expected to lead to a high variability in modifications. However, their negative counterparts for atypical values are quite effective. Applying “biped” to “human” has little effect on expectations about moving abilities, while applying “non-biped” has a crucial influence.

### 3 Experimental Data

The introduced extended modification model predicts that the occurrence and the direction of alteration by a modifier is determined by the existence of positive and negative constraints and that less typical modifiers result in larger changes. We contend that this is the rational way to handle the information one has about noun and

**Fig. 5** Mean ratings without constraint (NC) and with constraint (C) by modifier condition



modifier. We investigated whether people follow this strategy by a further analysis of the data from Connolly et al. (2007) and in an exploratory study we carried out.

### 3.1 *Constraint Influences in the Data of Connolly et al. (2007)*

If our extended modification model is accurate, it should be possible to find influences of constraints on the likelihood of modified sentences in the original data set by Connolly et al. (2007).<sup>4</sup> Our research group thus examined the original stimuli and agreed on constraints between modifiers and ascribed properties. A similar idea can be found in Jönsson and Hampton (2012), where the *subjects* were asked to justify higher or lower likelihood ratings of modified sentences. The main reasons given were pragmatic (e.g., the weirdness of the modified sentences), justifications by background knowledge about the modifier, or uncertainty about the modified noun.

We determined constraints for 5 B-modifiers, 14 C-modifiers and 21 D-modifiers. The mean ratings for constrained and unconstrained sentences by question type are shown in Fig. 5. The decrease in judged likelihood is much stronger for the constrained sentences. However, this result has to be interpreted keeping in mind that our post hoc analysis results in different sample sizes (e. g. 350 ratings for the unconstrained B-condition compared to 50 ratings for the B-constraint condition).

Since modifications do not necessarily decrease, but in some cases may increase the likelihood of a property (compare “Hamsters live in cages” and “Pet hamsters live in cages.”), it makes sense to look at the absolute values of the differences to the baseline conditions A, shown in Table 2. Here, the difference between constrained

<sup>4</sup>This analysis was made possible because the authors kindly provided us with their original data.



**Table 2** Mean absolute differences from the baseline condition A without and with constraint and in total

	A-B	A-C	A-D
No constraint	0.78	1.123	1.405
Constraint	1.46	2.3	2.319
Total	0.856	1.535	1.885

**Table 3** Results of post-hoc significance test (insignificant results shaded)

	A	B-NC	C-NC	D-NC	B-C	C-C
B-NC	0.010					
C-NC	0.000	0.007				
D-NC	0.000	0.000	0.996			
B-C	0.000	0.175	1.000	1.000		
C-C	0.000	0.000	0.002	0.151	0.728	
D-C	0.000	0.000	0.000	0.001	0.160	1.000

and unconstrained versions is even more obvious: the reduced likelihood is nearly twice as high for the constrained sentences. Furthermore, there is almost no difference between the simple (C) and double (D) modification, indicating that constraints have a stronger influence on the judged likelihood than modification.

The results of t-tests between all groups with Hochberg’s GT2 correction (for different sample sizes) are shown in Table 3. All groups differ significantly from the baseline condition A. The differences between constrained and unconstrained sentences are significant ( $p < 0.01$ ), except for the constrained B-condition, which is likely explained by its small sample size. The differences between the C- and D-conditions are not significant, neither for the constrained nor for the unconstrained version.<sup>5</sup> The results indicate that a more accurate grouping of the sentences would be between constrained and unconstrained modifications and neglecting the effect of double modification.

These analyses only allow for tentative conclusions because of the different sample sizes. But we can see a clear tendency in accordance with the predictions of the extended modification model: the change in likelihood ratings in the original data was shown to be much more distinguished for sentences in which the chosen modifiers constrain the assigned property.

<sup>5</sup>Jönsson and Hampton (2012, p. 98) also found insignificant differences between C- and D-stimuli in post-hoc pairwise comparisons.

## 3.2 Experiments

In order to test several of our empirical predictions we designed an *exploratory study* with few items and a comparatively small group of subjects. The described experiment served as a preparation of a larger study, reported elsewhere (Strößner and Schurz 2020). We tested several question types on four items.

### 3.2.1 Method

#### Participants

Subjects were 48 students of the Heinrich-Heine-Universität Düsseldorf, who were paid for participation.

#### Material

We used German translations of four items from Connolly et al. (2007). Two of them were previously judged to have no constraint between modifier and ascribed property by the members of our research group. For the third and fourth item, the modifiers were suspected to have a constraint on the typical property. This pre-experimental classification by the authors was used in order to look whether items with a suspected knowledge constraint behave differently. Previous studies by (Jönsson and Hampton 2012) have shown that subjects are often aware of subtle dependencies between modifier and property if they have to justify a lower likelihood rating for the modified sentence. It has been noted that these justifications could also be made up only after the rating task rather than really influencing it (cf. Gagné and Spalding 2014, p. 1290). Moreover, we were also interested to know whether knowledge constraints are purely subjective or intersubjective. If constraints are purely subjective, then there should be little differences between the items with constraint and the items without constraint. In addition to the preclassification by the authors, we also gathered relevance ratings from the subjects, which will be reported below. The double-modification was only tested for the two items with presumed non-relevant modifiers. The items are shown in list 1. The questions types are listed in list 2. Subjects gave ratings on the typicality and likelihood of the items (question type P and T) as well as on the typicality and likelihood of the modifier (question type PM and TM). The relevance rating was gathered with question type RM.

## 1. Lambs

- A Lämmer sind weiß. (Lambs are white.)
- B Flauschige Lämmer sind weiß. (Fluffy lambs are white.)
- C Norwegische Lämmer sind weiß. (Norwegian lambs are white.)
- D Langhaarige, norwegische Lämmer sind weiß. (Long-haired Norwegian lambs are white.)

## 2. Shirts

- A Hemden haben Knöpfe. (Shirts have buttons.)
- B Baumwollhemden haben Knöpfe. (Cotton shirts have buttons.)
- C Kratzige Hemden haben Knöpfe. (Itchy shirts have buttons.)
- D Kratzige Leinenhemden haben Knöpfe. (Itchy canvas shirts have buttons.)

## 3. Limousines

- A Limousinen sind lang. (Limousines are long.)
- B Teure Limousinen sind lang. (Expensive limousines are long.)
- C Preisgünstige Limousinen sind lang. (Inexpensive limousines are long.)

## 4. Sofas

- A Sofas stehen im Wohnzimmer. (Sofas are in living rooms.)
- B Bequeme Sofas stehen im Wohnzimmer. (Comfortable sofas are in living rooms.)
- C Unbequeme Sofas stehen im Wohnzimmer. (Uncomfortable sofas are in living rooms.)

**List 1** Items used in our experiment

- P Subjects rated the likelihood of the property for the unmodified and modified nouns.
- T Subjects rated the typicality of the property of the unmodified and modified nouns.
- PM Subjects rated the likelihood of the modifiers for the nouns.
- TM Subjects rated the typicality of the modifiers for the nouns.
- RM Subjects rated whether the modified attribute is relevant for the target attribute.

**List 2** Question types**Design**

In the first questionnaire, subjects were instructed to answer how typical they rate the default property, e.g. *being long*, for the modified and unmodified nouns, e.g. *limousines*, *expensive limousines* and *inexpensive limousines* (question type T). One group of 19 participants rated the items 1 and 3. Another group of 19 subjects rated items 2 and 4. Both groups also rated the typicality of the modifiers, e.g. *being*

*expensive* and *being inexpensive* for limousines (question type TM). In the second questionnaire, the subjects of both groups were asked to rate the likelihood of the same items (question type P and PM). The likelihood ratings and typicality ratings were gathered in two separate questionnaires but came from the same subjects. The unmodified and modified conditions as well as the rating of the modifiers were mingled but appeared on the same questionnaire. The participants thus saw their own answers and were potentially able to review and revise them. The last questionnaire contained relevance ratings (question type RM) for all items and modifiers. Subjects rated whether the modified attribute is relevant for the target attribute, e.g. whether the length of a limousine is related to its price. All judgements were given on a scale from 0 to 10. For the relevance question, subjects had the possibility to answer “I don’t know”.

### 3.2.2 Results

#### Typicality and Probability

In our model, probability plays a crucial role for defining typicality. We thus tested, whether the typicality ratings and the likelihood ratings are similar. This question is also important because even in a probabilistic approach, there are different notions of typicality: Schurz (2012) distinguishes *typicality in the wide sense* as probability in the category from *typicality in the narrow sense*, where a property has also to be improbable in sibling categories, i.e. highly discriminatory. This second criterion is what Rosch (1978) terms *cue validity*. For example, having a heart is only typical in the wide sense for birds. Having a beak is also typical in the narrow sense. Typicality in the wide sense justifies prediction of properties from known membership. Typicality in the narrow sense also allows to infer membership from known properties.

Table 4 shows the frequencies of the difference of all typicality ratings compared to the respective likelihood ratings. The typicality and likelihood ratings were very similar. In more than half of the pairs, they were even rated exactly the same.

**Table 4** Likelihood compared to typicality: cases and percentages

	+8	+7	+6	+5	+4	+3	+2	+1	0	-1	-2	-3	-4	-5
	2	1	3	8	14	24	37	64	355	77	33	10	5	10
%	0.3	0.2	0.5	1.2	2.2	3.7	5.8	10	55.2	12	5.1	1.6	0.8	1.6

Paired sample t-tests for the 24 pairs showed that only four pairs were significantly different.<sup>6</sup> The result strongly indicates that subjects preferred a wider notion of “typical” in the task. This supports our definitions of typicality in terms of probability.

### The status of the modifier

The data on the typicality and likelihood of the modifier in relation to the head noun allowed us to confirm that the B condition modifiers were also considered as typical by subjects in the German speaking community in comparison to the modifiers in the C and D condition. The mean values for B modifiers were clearly above 5, while the C modifiers were clearly below 5 in both, the typicality and the likelihood rating.<sup>7</sup>

### Comparison to Connolly et al. (2007)

Our main goal was to reproduce the modification effect in Connolly et al. (2007) with a possibility to distinguish between items with and without relevant constraints. For a better comparability, we converted our data to their 1–10 scale. Table 5 shows the descriptive statistics for likelihood and typicality ratings of the four items in comparison to theirs. The two tables show the means and the 0.95 confidence intervals for the probability question and the typicality rating. If the four items are considered together, the ratings resemble Connolly et al.’s result. As we already suspected, the data look quite different if the two relevant and the two non-relevant items are considered separately. The general loss under likelihood for the non-typical modifier is almost solely explained by the data for the relevant items *limousines* and *sofa*. The confidence intervals indicate that the differences from A to C (and also from B to C) are significant for the relevant items but not for the irrelevant ones.

### Relevance Correlations

We showed that the extent of the modification effect was predictable from our modifier relevance assumptions. But to what degree did our assumptions correspond to the the subjects’ ratings? And to what degree did their subjective relevance ratings correlate with their individually given likelihoods of the modified statements?

---

<sup>6</sup>Subjects rated having buttons to be more probable than typical for itchy shirts (1.211 [0.013, 2.406],  $p = 0, 048$ ), being long-haired more probable than typical for lambs (0.579 [0.129, 1.029],  $p = 0, 013$ ), being long more probable than typical for inexpensive limousines (1.221 [0.091, 2.330],  $p = 0, 036$ ) and being comfortable less probable than typical for sofas (−0, 421, [−0, 825, −0, 017],  $p = 0, 042$ ). Brackets give mean value with 0.95 confidence intervals.

<sup>7</sup>The lowest mean value of a B modifier was 5.92 [5.33, 6.50] in the likelihood rating of *cotton* for *shirt*. The highest mean value of a C modifier was 3.45 [2.70, 4.28] in the likelihood rating of *itchy* for the same item.

**Table 5** Modification effect in comparison to Connolly et al. (2007)

	Shirts, Lamb	Limousines, Sofas	All	Connolly et al. (2007)
(a) Likelihood ratings				
N	37	38	75	400
A	8.30 [7.69, 8.91]	8.37 [7.82, 8.91]	8.33 [7.93, 8.73]	8.38 [8.20, 8.56]
B	7.76 [7.14, 8.39]	7.94 [7.36, 8.52]	7.85 [7.44, 8.27]	7.72 [7.51, 7.93]
C	7.49 [6.69, 8.23]	4.62 [3.75, 5.50]	6.04 [5.37, 6.71]	6.89 [6.66, 7.12]
D	7.42 [6.66, 8.18]	–	–	6.50 [6.27, 6.73]
(b) Typicality ratings				
N	38	38	76	400
A	8.15 [7.54, 8.76]	8.27 [7.65, 8.89]	8.21 [7.79, 8.84]	8.38 [8.20, 8.56]
B	7.77 [7.04, 8.51]	7.80 [7.13, 8.46]	7.79 [7.30, 8.27]	7.72 [7.51, 7.93]
C	7.13 [6.38, 7.89]	4.10 [3.20, 5.00]	5.62 [4.95, 6.29]	6.89 [6.66, 7.12]
D	7.13 [6.36, 7.91]	–	–	6.50 [6.27, 6.73]

**Table 6** Mean judged relevance of non-typical modifiers with 0.95 confidence intervals

Limousines	Sofas	Shirt	Lambs	Lambs (excl)
N = 38	N = 38	N = 38	N = 38	N = 32
3.53	3.37	0.32	3.00	3.56
[2.51, 4.55]	[2.28, 4.48]	[0.07, 0.56]	[1.92, 4.08]	[2.38, 4.75]

The non-typical modifiers were judged to be more relevant for the items *limousines* and *sofas* than for *shirts*. However, for *lambs* people judged origin to be more relevant for the colour than we expected. The item also stood out insofar as many people suspended judgement on this relevance question, while no subject answered “I don’t know” in the relevance rating of any other non-typical modification.<sup>8</sup> Table 6 shows the relevance judgements for the nontypical modifiers. “I don’t know” answers were treated as “0” in the column “Lambs” and excluded in “Lambs (excl)”.

Finally we wanted to know whether the differences in the subjects’ likelihood ratings are related to their individual relevance ratings. We tested this hypothesis for the non-typical modification. First, we calculated the individual modifier effect by subtracting the judgement of the unmodified condition A from the judgement in the modified condition, i.e. C-A.<sup>9</sup> By that means we determined the modification effects for each individual and each item. These values were correlated with the relevance

<sup>8</sup>The differences are probably explained by our research group considering cross value dependencies while the subjects were only confronted with attributes. They might have regarded general evolutionary tendencies that living environments influence appearance, which we did not consider because they are not important for this particular values. We came to the conclusion that it is important to ask for the particular values.

<sup>9</sup>This was possible because we used a within-design.

**Table 7** Kendall's  $\tau$  correlation of relevance and loss in the likelihood rating

Limousines	Sofas	Shirts	Lambs
N = 19	N = 19	N = 19	N = 19
-0.37, $p = 0.046$	-0.38, $p = 0.04$	-0.69, $p < 0.01$	-0.43, $p = 0.03$

ratings of the 19 subjects. This correlation reflects the influence of subjective constraints. Thus we tested whether individually higher or lower modification effects come with individually higher or lower relevance assumptions. It turned out that the status of the modifier was correlated to the modification effect. Kendall's  $\tau$  revealed that subjects with larger decreases in the non-typical modification tended to find the modifier relevant. The correlation and significance is given in Table 7. The correlation is moderate for the items *lamb*, *limousine*, *sofa* and even high for *shirts*, which had a low intersubjective relevance score.<sup>10</sup>

## 4 Conclusion

In this paper, we proposed an extended modification model with constraints. An exploratory study with four of the items used by Connolly et al. (2007) revealed the following tendencies, which largely accord to our assumptions:

- Typicality and likelihood rating  
Likelihood ratings are very similar to typicality ratings. This supports probabilistic approaches to typicality.
- Typical modification  
As already suspected from previous studies, typical modifiers lead to a smaller loss in modification than non-typical ones.
- Valid prior classification for constraints  
The items we suspected to have negative constraints were drastically affected by non-typical modification while the difference to typical modification was negligible for items without constraints.

The prior predictability of loss by non-typical modification has never been investigated so far. However, Connolly et al. (2007) already suspected that the distinctiveness of modification effects is predictable, asserting that “adding *purple* to apple is sure to diminish one’s confidence about its edibility more than adding *ripe* and less than adding *Martian*” (Connolly et al. 2007, p. 14). They take this to be an argument against prototype compositionality. Though prototype compositionality is not

<sup>10</sup>In a later larger study with more item, reported in Ströbner and Schurz (2020), we were not able to confirm such high correlations between individual modification effects and individual relevance assumptions. However, we were able to confirm that the mean relevance score for a modification is correlated with its mean modification effect.

as straight forward as composing analytic meanings, we disagree with the conclusion that prototypes are not compositional at all. People systematically attribute typical properties to subcategories, even if they are built with meaningless words, as noted by Gagné and Spalding (2014). Their “different but similar” approach, however, would predict that all modifications have roughly the same effect. This is, however, not the case. Our model predicts differences for modifications without constraints and modifications with relevant knowledge constraints. Doubters of prototype theory could argue that the extended modification model includes background beliefs and is thus not about semantics but about belief revision. This argument depends on a very narrow view of compositionality. Understood in the sense of Hampton and Martin (2012) as a process that is not only driven by strict logical intersection but also by common-sense knowledge, the enriched modification model is a model of prototype compositionality.

**Acknowledgements** The research reported in this article was conducted in the project D01 (Frame representation of prototype concepts and prototype-based reasoning) of the SFB991, a collaborative research center that was generously funded by the German Research Foundation (Deutsche Forschungsgemeinschaft). We are grateful to Andrew Connolly for sharing data from his research with Jerry Fodor, Lila Gleitman and Henry Gleitman. Paul Thorn, Sebastian Löbner and an anonymous reviewer provided helpful comments on our manuscript.

## References

- Barsalou, L. W. (1992). Frames, concepts, and fields. In A. Lehrer & E. F. Kittay (Eds.), *Frames, fields, and contrasts*, pp. 21–74. Lawrence Erlbaum Associates Publishers, Hillsdale, New Jersey.
- Connolly, A. C., Fodor, J. A., Gleitman, L. R., & Gleitman, H. (2007). Why stereotypes don’t even make good defaults. *Cognition*, *103*(1), 1–22.
- Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*(2), 163.
- Fodor, J., & Lepore, E. (1996). The red herring and the pet fish: why concepts still can’t be prototypes. *Cognition*, *58*(2), 253–270.
- Gagné, C. L., & Spalding, T. L. (2011). Inferential processing and meta-knowledge as the bases for property inclusion in combined concepts. *Journal of Memory and Language*, *65*(2), 176–192.
- Gagné, C. L. & Spalding, T. L. (2014). Subcategorisation, not uncertainty, drives the modification effect. *Language, Cognition and Neuroscience*, *29*(10), 1283–1294.
- Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, *15*(1), 55–71.
- Hampton, J. A. & Martin, L. J. (2012). Typicality and compositionality: The logic of combining vague concepts. In W. Hinzen, E. Machery, & Werning M. (Eds.), *The Oxford handbook of compositionality*, pp. 385–402. Oxford University Press, Oxford.
- Hampton, J. A., Passanisi, A., & Jönsson, M. L. (2011). The modifier effect and property mutability. *Journal of Memory and Language*, *64*(3), 233–248.
- Jönsson, M. L., & Hampton, J. A. (2012). The modifier effect in within-category induction: Default inheritance in complex noun phrases. *Language and Cognitive Processes*, *27*(1), 90–116.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann.



- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization*, Chap. 2, pp. 27–48. Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology*, 8(3), 382–439.
- Schurz, G. (2012). Prototypes and their composition from an evolutionary point of view. In W. Hinzen, E. Machery, & M. Werning (Eds.), *The Oxford handbook of compositionality*, pp. 530–553. Oxford: Oxford University Press.
- Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive science*, 12(4), 485–527.
- Strößner, C., & Schurz, G. (2020). The role of reasoning and pragmatics in the modifier effect. *Cognitive Science*, 44(2). <https://doi.org/10.1111/cogs.12815>.

**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.

