# Concepts in context: Evidence from a feature-norming study

**Diego Frassinelli (d.frassinelli@sms.ed.ac.uk)**
Institute for Language, Cognition and Computation
School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB, UK

**Alessandro Lenci (alessandro.lenci@ling.unipi.it)**
Dipartimento di Linguistica "T. Bolelli", Via Santa Maria 36
56126 Pisa, Italy

**Abstract**

Concepts are typically conceived as context-free knowledge structures. Recently, a different view has emerged according to which subjects produce situation-specific conceptualizations, thereby raising important questions about the level of contextual dependency in conceptual representations. In this paper, we present a feature-norming study in which subjects are asked to generate properties of concepts presented in context. Collected data are analysed to investigate the actual amount of conceptual variation induced by contexts and the effect of context modality.

**Keywords:** Semantic feature norms; property generation; context.

## Concepts and contexts

Both in classical and in post-classical models, concepts have been conceived as substantially context-free knowledge structures. Regardless of the particular theory (e.g. exemplar, prototype, and connectionist), it is generally assumed that concepts result from abstracting critical information about an entity *per se* (such as shape, colour, etc.), leaving behind background situations (i.e. the contexts) in which these entities are experienced. Concepts thus become invariant to different contexts of use. Accordingly, the same representation of an apple is used both when categorizing an entity on a tree, and when categorizing the same entity in a supermarket.

Recently, this view has been overtly criticized. For instance, Yeh and Barsalou (2006) argue that concepts not only contain a large array of situational information about the physical settings, events, and subjective perspectives of agents, but they also produce different conceptualizations in different contexts. For instance, the supermarket situation would activate context-specific information concerning an apple, different from that activated by a different context, such a tree in a garden. These two claims directly follow from the perceptual simulation model adopted by the authors, but more in general they raise important questions about the level of contextual dependency in our conceptual representations. Wu and Barsalou (2009) used a property generation task to investigate the situated nature of concepts, and reported that approximately 26% of the features produced by subjects were indeed situation-related. Subjects generated properties (semantic feature norms) provide interesting evidence about conceptual representations, but one intrinsic limit of the study in Wu and Barsalou (2009) is that stimuli were presented *out of context*, as it is customary in semantic norming. This way, it becomes impossible to address and test the more specific and crucial issue concerning the relation between concepts and context, that is the actual effect of the context in modulating and biasing conceptual representations.

In this paper, we present a feature-norming study in which subjects are asked to generate properties of concepts presented *in context*. To the best of our knowledge this is the first property generation task with this design. While we do not commit ourselves to any specific model of conceptual representation, collected data allow us to address directly three key issues concerning the effects of different contexts on concepts: i.) the actual amount of conceptual variation induced by contexts, and ii.) the property types that are more subject to contextual variation, and iii.) the effect of the context modality. In particular, we will investigate the effect of both linguistic contexts (i.e. a sentence in which the context noun appears) and extralinguistic contexts (i.e. an image of a situation in which an entity can be experienced).

## Semantic Feature Norms

Nowadays there is a strong consensus on the fact that it is possible to describe the internal structure of a concept in terms of a set of semantic properties (Garrard, Lambon Ralph, Hodges, & Patterson, 2001; Baroni & Lenci, 2008). A traditional way to access and study the structure of conceptual knowledge is the use of semantic features norms. These are lists of properties that participants produce describing and defining a specific concept; moreover they include several measures and statistics calculated according to feature production frequencies.

As suggested by McRae and colleagues (McRae, Cree, Seidenberg, & McNorgan, 2005) these lists do not provide a static and definitive representation of concepts, however, they are the most direct way to study the dynamics associated with the online process that takes place when subjects have to process a specific concept.

Different researchers used these lists to investigate various aspects of human cognition. They have been used to test the psychological validity of cognitive theories (Wu & Barsalou, 2009), and as stimuli for different experiments such as semantic similarity (McRae, Sa, & Seidenberg, 1997) and property verification tasks (Cree, McNorgan, & McRae, 2006).

One of the most widely used norms is the collection of McRae et al. (2005). This is the largest set of semantic properties freely available: it includes properties for 541 living and non-living concepts. Another smaller example is represented by the collection of Vinson and Vigliocco (2008). In this case, the authors extended their analysis to the domain of actions and events. They collected norms for 167 living and non-living objects and for 287 events and actions.

The pros of these collections are relatively straightforward; however they exhibit also different limitations (McRae et al., 2005). The process of collection, normalization and classification is extremely long and expensive. Moreover, the linguistic nature of the task favours the information which is easily verbalized, penalising spatial and temporal relations between entities. During the classification phase, the annotators have to reinterpret the intents of the subjects and cannot always preserve the original information. Finally, as we said above, all existing norms were collected by presenting words in isolation and not associated with a specific context. In this work, we will focus our attention on this last feature.

## Collecting context-sensitive feature norms

The main goal of this work is to describe the collection of semantic feature norms for 8 concrete concepts and to analyse the effects that contextual variability exerts on the number and types of properties produced.[1]

### Design

The collection was performed on-line using a website interface.

**Stimuli** The 8 normed concepts correspond to the following English nouns: *apple*, *banana*, *bear*, *horse*, *bike*, *car*, *hammer*, and *knife*. The nouns were sampled in order to have an equal number of animate and inanimate concepts belonging to the semantic classes traditionally used in these studies, that are fruits, animals, vehicles, and tools.

For each concept, we identified two alternative situations frequently associated with the correspondent object. We downloaded from the Web 16 colour pictures depicting the two contexts for each concept and we downsized them (288*320 pixels). The pictures do not include the target object unless it is strictly necessary for the correct interpretation of the context (e.g. a showroom without some cars inside would not be identifiable). This way, participants are not biased in their descriptions by a specific instance of the concept appearing in the picture. A native English speaker produced 16 sentences describing the context depicted in the correspondent picture. Unlike the visual contexts, the sentences include the target concept noun (written in capital letters).

For every trial, the target concept noun (in capitals) appears on the top of the screen and is followed by 10 blank lines that participants have to fill in with concept properties. In the case of the linguistic context, the sentence containing the target word appears instead of the target word. For the visual context, the picture appears on the left of the blank lines. Figure 1 shows the visual and linguistic contexts for *apple*.

(a) the APPLE is ripening on the tree

(b) A man is buying an APPLE at the market

Figure 1: *apple*: visual and linguistic contexts.

**Procedure** The entire experiment included 40 different combinations of the 8 concepts and the 5 different context types (2 visual, 2 linguistics, and 1 no context). 125 lists of 8 items each were created: the distribution of the items across lists was based on a Latin Square design, ensuring that every list comprised only one occurrence of each concept and one of each specific context. Every list included the 8 concepts as follow: one or two out of context, three or four in a visual and linguistic context respectively. The order of the trials was semi-random: all the out of context trials appeared before the linguistic ones and those before the visual ones. In this way, the complexity of the stimulus increased during the experiment.

125 native English speakers recruited online performed the experiment. Each participant saw the 8 items of one of the 125 lists; in this way, every concept and context were seen only once. The experiment started after written instructions of the task and 3 examples. The task required to read the word (or the sentence) at the top of the screen and to produce a maximum of 10 properties per concept describing different aspects of it. The instructions clearly stated that the aim of the experiment was to study how people process the meaning of words; subjects were not instructed to take contextual variability into account. Moreover, we provided a list of possible qualities of the concepts to take into account during the actual experiment: colours (e.g. CHERRY *red*)[2], tastes (e.g. ICE CREAM *good*), shapes (e.g. BALL *round*), functions (TRAIN *transportation*), typical locations (SHARK *ocean*), emotions (CHRISTMAS *excitement*), evaluations (SOUP *hate*), etc. We did not set a maximum amount of time for a single trial however, on average, the entire experiment lasted about 15 minutes.

[1] The collection is freely available at `http://sesia.humnet.unipi.it/norms`

[2] In this work, we use caps to indicate CONCEPTS and italics to indicate the *properties produced*.

### Post Processing

**Data Codification** After collecting the features, a process of data filtering and normalization was carried out. We identified all the synonym properties and we normalized them to the same feature (e.g., "bike" and "bicycle" were coded as *a bike*). Coordinate or disjunctive features providing more than one piece of information were split into separate tokens. For example, "is red, green or yellow" became *is red*, *is green*, and *is yellow*. We removed all the quantifiers (e.g. "can be", "generally", "usually") and other materials not relevant to the analysis (e.g. miscellanea, incoherences, and free associations). Finally, one of the authors coded the resulting properties according to a specific set of patterns (e.g. BEAR *beh_eats_honey* codes "to eat honey" as a prototypical behaviour (beh) of bears) and classified them according to the scheme described below.

**Coding Scheme** The properties were classified according to a partially simplified version of the coding scheme proposed by Wu and Barsalou (2009). The scheme includes 24 property types grouped into 4 main categories:

- *Taxonomic properties* (TAX): properties describing taxonomic relations (hypernyms, hyponyms, synonyms, and coordinates).
- *Entity properties* (ENT): properties describing the entity per se (e.g. internal and external properties and elements, prototypical behaviours).
- *Situation properties* (SIT): properties associated with the contextual background (e.g. locations, time, participants, functions).
- *Introspective properties* (INT): properties describing feelings and mental states (e.g. evaluations, contingencies).

## Results

Participants produced 6922 properties in total: 3619 entity properties, 2025 situation properties, 644 introspective properties, and 634 taxonomic properties. Table 1 reports the average number of features (and Standard Error) produced by every subject for each item and grouped by property class and modality. There are no noticeable differences among modalities (visual, linguistic, and no context): this suggests that different contexts are not exerting a significant effect on the number of properties generated. The differences arise analysing the property classes. The properties describing different aspects of the target concept (ENT) are the most frequent (52% of the total). Properties providing contextual information related to the target concept (SIT) are produced the 29% of the time (interestingly, this figure is close to the one reported by Wu and Barsalou (2009)). Properties describing mental states (INT) and taxonomic relations (TAX) are less frequent (around 9%).

Table 1: Average (AVG) and Standard Error (SE) of the number of features produced by each subject for each concept grouped according to broad property class and modality.

| Class | No Context | | Visual | | Linguistic | |
|---|---|---|---|---|---|---|
| | AVG | SE | AVG | SE | AVG | SE |
| TAX | 1.20 | 0.09 | 1.34 | 0.15 | 1.24 | 0.11 |
| ENT | 3.82 | 0.40 | 3.64 | 0.36 | 3.86 | 0.41 |
| SIT | 2.34 | 0.26 | 2.66 | 0.31 | 2.15 | 0.25 |
| INT | 1.61 | 0.19 | 1.59 | 0.19 | 1.64 | 0.19 |

### Analysis

**Model** We analysed the data adopting the framework of the linear-mixed effects models with a Poisson linking function (Baayen, Davidson, & Bates, 2008). The dependent variable was the property frequency. Table 2 presents the coefficients and p-values of the mixed model. To investigate the effects exerted by contextual variability we included two factors in contrast coding: the factor `Modality` for the effects of visual (+.5) and linguistic context (-.5), and the factor `Context` for the effects produced in the out-of-context (-.5) and in the in-context (+.25) conditions. We also analysed the effects associated with the type of feature produced: the factor `Property` compared the object related properties such as entity and taxonomic properties (-.25) and the context related properties such as situation and introspective properties ( +.25); the factor `ObjectProp` coded the effects of taxonomic (-.5) and entity (+.5) properties; and the factor `SituationProp` the effects of situation (-.5) and introspective (+.5) properties. The random effects were `Subject` and `Concept`, which were intercepts in the model. We also included random slopes for all the main effects (Modality, Context, Property, ObjectProp, and SituationProp).

Table 2: The coefficients for the linear-mixed effects model.

| Predictor | Coefficient | Signif. |
|---|---|---|
| (Intercept) | 0.66 | *** |
| Property | -0.28 | |
| ObjectProp | 1.10 | *** |
| SituationProp | -0.43 | *** |
| Context | 0.01 | |
| Modality | 0.04 | |
| Context:Property | -0.09 | |
| Context:Object | -0.17 | |
| Context:Situation | -0.06 | |
| Modality:Property | 0.14 | |
| Modality:Object | -0.13 | |
| Modality:Situation | -0.29 | ** |

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

**Property Types** The factor `Property` compares the overall mean of entity and taxonomic properties (properties more associated with the object) with the overall mean of situation and introspective properties (properties that are more associated with the context). We find a slightly significant effect ($p < .1$) in favour of the first group of properties: object related properties are produced more frequently than context related properties. The factor `ObjectProp` shows a significant positive main effect for the entity properties compared to the taxonomic ones. The factor `SituationProp` reveals a positive effect for situation properties compared to the introspective ones.

**Context and Modality** As expected, there is no significant main effect for `Context` and `Modality`. We only find a not significant effect associated with the visual context. Participants produced almost the same number of properties independently of the presence and type of contextual information they are exposed to.

**Interactions** The interactions reported in table 2 do not reveal significant effects. The only significant effect is described by the interaction `Modality:SituationProp` where situation properties are positively biased by visual contexts, while introspective properties are more biased by linguistic properties.

## Qualitative Analysis of Feature Types

In this section we present a qualitative analysis of the data to determine the main effects exerted by contextual information on specific property types.

For each concept, we divided the features produced in both contexts from those associated only with a specific context. We did the same procedure for visual and linguistic contexts independently. The aim of this analysis is to determine which property types are more dependent on a specific context (context dependent) and which are produced in both contexts (context independent). We are interested in a general evaluation of this effect without taking into account interconceptual variability: for this reason we combined the results obtained for each concept. After a preliminary analysis, we discovered that property types show an almost coherent trend both in a visual and linguistic context. We analysed the effects exerted by the two modalities using a linear model. We did not find a significant difference between visual and linguistic modality as main effect ($\beta_{Visual} = 3.90, p = 0.94$) and also in interaction with the context_dependent vs context_independent variable ($\beta_{Visual:ContextInd} = 4.54, p = 0.88$). For that reason, table 3 reports the results from a general point of view, without modality distinction. If there is a difference between the two modalities we will discuss it separately. We report the percentage of context dependent and context independent properties out of the total number of properties of the same type (e.g. the 92% of hypernyms are context independent). We also present in bold the percentage for the entire class at the end of each group of properties (e.g. the 80% of taxonomic properties are context independent). In brackets there is the number of properties of each type out of the total number of properties in the same class (e.g. the 78% of taxonomic properties are superordinates).

Table 3: Percentage of the features that are context dependent (associated with only one context) and context independent (associated with both contexts).

| Property | Dependent | Independent |
|---|---|---|
| C-super (.78) | 0.08 | 0.92 |
| C-subord (.19) | 0.57 | 0.43 |
| C-coord (.02) | 1 | 0 |
| C-syn (.02) | 0.5 | 0.5 |
| **Taxonomic** | **0.2** | **0.8** |
| E-exsurf (.27) | 0.15 | 0.85 |
| E-excomp (.24) | 0.12 | 0.88 |
| E-sys (.21) | 0.28 | 0.72 |
| E-beh (.07) | 0.39 | 0.61 |
| E-incomp (.06) | 0.24 | 0.76 |
| E-insurf (.06) | 0.17 | 0.83 |
| E-mat (.06) | 0.07 | 0.93 |
| E-quant (.02) | 0.53 | 0.47 |
| E-whole (.01) | 0.21 | 0.79 |
| **Entity** | **0.2** | **0.8** |
| I-cont (.68) | 0.62 | 0.38 |
| I-eval (.30) | 0.38 | 0.62 |
| I-emot (.02) | 0.5 | 0.5 |
| **Introspective** | **0.54** | **0.46** |
| S-func (.47) | 0.25 | 0.75 |
| S-assoc (.15) | 0.58 | 0.42 |
| S-loc (.15) | 0.44 | 0.56 |
| S-action (.08) | 0.33 | 0.67 |
| S-particip (.08) | 0.48 | 0.52 |
| S-origin (.06) | 0.1 | 0.9 |
| S-time (.01) | 0.59 | 0.41 |
| S-socart (<.01) | 1 | 0 |
| **Situation** | **0.35** | **0.65** |

**Taxonomic Properties** Taxonomic properties describe highly stable relations among concepts. As expected, the 80% of these properties are equally produced in different contexts. The hypernyms (C-super, e.g. CAR *a vehicle*) are the 78% of the entire taxonomic class. These properties are represented by a small number of highly frequent feature types (in total only 22 for the visual and linguistic modalities) describing associations strictly language related. Hyponyms (C-subord, e.g. APPLE *Granny Smith*) include a high number of infrequent property types (in total 48 subordinates) and are more context dependent. This can be expected, given that each concept is associated with many hyponyms, which in turn might become differently prominent depending on the context. Coordinates (C-coord, e.g. APPLE *a pear*) and synonyms (C-syn, e.g. CAR *an automobile*) are only a small group of properties.

**Entity properties** The trend of this class is consistent: all the properties describing objects' qualities are not sensitive to contextual variability (the 80% of the total). The only exception is represented by those properties describing frequency or intensity (E-quant, e.g. APPLE *different varieties*); however this group includes a very small number of features and it is valid only for the linguistic modality (65% of context dependent properties).

**Introspective properties** The most substantial group among introspective properties is represented by contingency properties (I-cont, e.g. APPLE *is good with cinnamon*). These properties describe the "common sense knowledge" associated with a specific object in specific conditions. For this reason, it is not surprising to see a strong contextual effect. On the other hand, evaluations about the object (I-eval, e.g. APPLE *is delicious*) are less context dependent: participants have a personal opinion about every object that is unlikely to change in different situations. Emotions (I-emot, e.g. BEAR *is scary*) are very few cases.

**Situation Properties** The behaviour of this group of data is more various, given also the high heterogeneity of the properties in this class. Some properties are intrinsically related to an entity, and, therefore, less variable across situations: for instance, typical functions (S-func, e.g. CAR *used for transportation*), actions (S-action, e.g. APPLE *used by cooking*), origins (S-origin, e.g. APPLE *grows on trees*) and locations (S-loc, e.g. BANANA *grows in tropical climates*). Instead, other property types are more context-related, and, therefore, subject to stronger cross-situation variation, such as associations (S-assoc, e.g. CAR *associated with speed*). Participants (S-particip, e.g. BANANA *eaten by monkeys*) are almost equally present in both sets.

### Analysis of Feature Density

The last analysis we carry on compares the distribution of specific features in terms of feature density: how many subjects produce the same feature for the same stimulus. We analysed the features divided into the context dependent and context independent sets. In the context dependent set the 85% of specific properties are produced by only one person, the 11% by two different subjects. The remaining 4% is shared by properties produced from 3 to 10 people. On the other hand, in the context independent set we have the 21% of features produced only by 2 subjects (the minimum value for having an overlap). The maximum number of subjects that produced the same feature is 47 (KNIFE *is sharp*).

## General Discussion

The experiment described in this work was aimed to test the effects exerted by contextual variability on the production of semantic properties by human beings. We gave both quantitative and qualitative evidence of these effects. Neither modality nor context variability has significant effects on the number of features produced. People list almost the same number of properties in different contexts. However, these properties are not equally distributed and the differences among them are statistically significant. As already emerged in the literature, subjects produce more entity properties than taxonomic ones and more situated properties than introspective ones. It is interesting to note that merging together the properties more object related (entity and taxonomic) and the properties more context related (introspective and situation) the difference decreases considerably with only a slightly significant effect in favour of the first group. This suggests that people are including in their dynamic representation of concepts both information describing the object per se but also almost the same amount of background information. To gather more evidence, we performed also a qualitative analysis. In this case, we extracted from the collection only the properties produced in context and we identified those occurring with both contexts and those associated with only one. The results are straightforward for the taxonomic and entity properties: almost the 80% of all the properties classified in this way are produced in both contexts. We assist to an opposite effect when we move to the introspective properties. More complex is the dynamic of situation properties: some of them are more related to the object, some others to the context.

These results suggest that the context sensitivity of concepts is strongly limited to certain property types. A possible explanation can be found in Barsalou (1982). In this work, the author suggests the existence of two different kinds of properties: *context independent properties* strictly associated with the object *per se*, and *context dependent properties* associated with the specific context in which the word appears. Our data point in the same direction. It is possible to identify a large group of "core" properties that are not biased by contextual variability (in particular entity and taxonomic properties, as expected) and a smaller group of more dynamic properties produced less frequently and only associated with specific contexts (introspective properties, and partially situation properties).

## Acknowledgements

The experiment was funded by the Computational Linguistics Lab (University of Pisa) and the European Research Council under award number 203427 "Synchronous Linguistic and Visual Processing". We thank Larry Barsalou, Marco Baroni, Frank Keller and Moreno I. Coco for their help and suggestions.

## References


Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baroni, M., & Lenci, A. (2008). Concepts and properties in word spaces. *Italian Journal of Linguistics*, *20*(1), 53–86.

Barsalou, L. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition*, *10*(1), 82–93.

Cree, G. S., McNorgan, C., & McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *Learning, Memory, and Cognition*, *32*(4), 643–658.

Garrard, P., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2001). Prototypicality, distinctiveness, and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts. *Cognitive Neuropsychology*, *18*(2), 125–174.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–59.

McRae, K., Sa, V. R. de, & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology*, *126*(2), 99–130.

Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*(1), 183–190.

Wu, L., & Barsalou, L. (2009). Perceptual simulation in conceptual combination: evidence from property generation. *Acta Psychologica*, *132*(2), 173–89.

Yeh, W., & Barsalou, L. (2006). The situated nature of concepts. *American Journal of Psychology*, *119*(3), 349–84.