PIA BÖRGERDING, MARIE-CHRISTINE BENEN, AND ALEXANDER BERGS

Expecting the Unexpected? Predictive Coding, Pattern Recognition, and Surprise in Narratives

Introduction

One of the aesthetic pleasures of reading literature is the discovery of the unexpected, the new, the surprising. The surprise may be a twist in the plot (as, for example, when we realize that most of the story in Ian McEwan's *Atonement* (2001) was made up by the main character Briony), an unexpected or highly unlikely event or fact (as, for example, when we get to learn that the creature in Mary Shelley's *Frankenstein* [1818] is not a mindless monster, but a rational being with thoughts and feelings), or some unusual stylistic device (such as the drawing of the plot as a "tolerable straight line" in chapter XI of Lawrence Sterne's *Tristram Shandy* [1759-67]). The surprise may even lie in violated genre expectations, as, for example, when we pick up Danielewski's *House of Leaves* (2000) and expect a 'normal' novel.

Surprising and unexpected phenomena in literature have received attention from scholars since the early 20th century. Especially the Russian formalists were driven by the question of what makes a piece of work literature and argued for defamiliarization as a key phenomenon in literature, i.e., the presentation of familiar concepts in an unfamiliar way, which is achieved by deviating from linguistic norms (cf. Shklovsky [1916] 2017). Even though the assumption that deviation is what makes literature special is supported by various scholars, the main weakness of the formalists' approach is that they did not define the norm that literary language deviates from (cf. Cook 1994, 139; Strasen 2008, 248) as well as focusing solely on the linguistic properties of the text itself instead of taking into account its context and content (cf. Jeffries 2001, 326).

One attempt at solving this problem involves moving away from purely textual features towards combining defamiliarization and schema theory. Schema refreshment, as in challenging existing schemata (on the levels of language, text, and world knowledge) and their constellations, is seen as a crucial feature of literary texts since they provide the opportunity of experimenting with new patterns without risking failed communication (cf. Cook 1994). Other researchers follow Cook's approach but suggest some alterations: Semino (1995), for instance, points out that not only complete schema change but also schema reinforcement as well as connections between usually unrelated schemata add to the effects of literary texts. Jeffries does not deny that literary texts have the potential of changing or reinforcing schemata, but claims that literature does not necessarily need to provide "a challenge to existing schemata but the thrill of recognition" (2001, 334).

Other lines of research examine the emotional effect that defamiliarized texts have on the reader. Empirical studies show that the use of deviating linguistic structures, such as metaphors, alliterations or others, attracts the attention of readers but simultaneously slows down the processing speed (cf. Miall and Kuiken 1994). Foregrounded passages also have an effect on the manner of processing in readers –

While we agree that not only textual but also contextual features need to be considered in researching the effects of literature, we do not assume that literary texts inherently hold a special status in language. Due to the Construction Grammar claim of constructions being form-meaning pairings, an occurring linguistic form is accompanied by its associated meaning consisting of semantic, pragmatic, and contextual information (cf. Bergs and Diewald 2008; Bybee 2010; Hoffmann and Bergs 2015). However, literature does not (only) consist of special constructions, it is rather often characterized by unusual combinations of constructions (cf. the papers in Hoffmann 2018), occurrences of constructions where the reader would not expect them, or, as in the case of our experiment, variations in already established constructions in order to surprise the reader.

But how is this kind of surprise processed by actual flesh-and-blood readers? How do they deal with the unexpected? How many surprises can they take before they deem the input nonsense and stop processing it? This paper presents an exemplary empirical case study that probes into cognizers’ abilities and willingness to deal with unexpected input. It utilizes the model of probabilistic thinking, Bayesian inference, and predictive coding in the processing of (fictional) narratives (cf. Kukkonen in press; 2014). However, in doing so, we acknowledge the fact that predictive coding as such is still a new and critically debated neuropsychological model that has only been tested for very clearly defined data sets and domains (such as visual perception). As a consequence, we will primarily base our arguments not on empirical studies of predictive coding for full narrative texts (which would be a highly difficult task at this point), but on a more limited and clearly defined data set, the processing of idiomatic chunks or constructions, and variations thereof.

The reasoning is as follows: based on empirical data on the processing of these idiomatic constructions, we will show that predictive coding, or probabilistic cognition, is at work in the processing of (complex) linguistic material. Hearers predict (and expect) certain linguistic structures as input, depending on both co(n)text and general principles, such as embodied cognition. Unexpected input requires more processing, which does not necessarily lead to unacceptable or unparsable results, but rather triggers other cognitive processes that need to deal with the prediction error (in an effort after meaning; see Bartlett 1932), including an adjustment of the predictive system (in a Bayesian sense). What we show here in this study is a miniature sandbox experiment for this kind of model, but we expect that similar processing principles are at work in the understanding of more complex linguistic input, such as narrative texts. This is interesting in particular for what has been described as ‘unnatural narratives’ (see Alber and Heinze 2011; Alber et al. 2010; Richardson 2015), where readers encounter impossible and hence probably also unlikely and unexpected facts, events, and twists of plot. So how do actual readers react to and process the flying island of Laputa in Swift’s Gulliver’s Travels (1726)? How do they deal with the use of different font colors in Danielewski’s House of Leaves?
Predictive Coding and Bayesian Probabilities

Predictive coding models claim that the brain (or mind) is constantly working with more or less abstract (Bayesian) predictions about (future) sensory input. If, for example, we encounter the sequence of three individual sounds from an actual chord (say, C, E, and G) we intuitively predict that next time we hear C followed by E, there will be a third sound, namely G, to follow these two. If we encounter another chord, say, C, E, and G#, we might even form the more abstract prediction that sounds will always come in three and that they will somehow harmonize (in the form of a chord). In a Bayesian model, these predictions can get stronger over time through feedback loops with actual sensory input, in that sensory input from the lower levels of cognition is suppressed. Only if there is a difference between prediction and actual sensory input and this difference reaches a certain threshold will we see an update of the higher-level predictive system in order to accommodate the non-matching input into a better predictive model. In other words, only new information that passes the prediction error threshold is relevant here. The element of surprise is triggered once our predictions or expectations of sensory input are not met and our predictive systems need to be updated.

This model has been primarily developed and tested for 'simple' sensory perception and vision in particular (see, e.g., Rao and Ballard 1999). Barrett (2016; 2015) also applies the model in her analysis of interoceptive cognition and emotions; Seth (2013) shows that it can also be utilized in the study of how we understand other human agents and their internal states, including emotions.

Predictive coding and Bayesian inferences appear to be very similar, even compatible models; yet, the relationship between the two is still difficult to determine (cf. Kwisthout and van Rooij 2013; Fink and Zednik 2017; Aitchison and Lengyel 2017). Aitchison and Lengyel, for example, argue that "predictive coding is an algorithmic/representational motif that can serve several different computational goals of which Bayesian inference is but one. Conversely, while Bayesian inference can utilize predictive coding, it can also be realized by a variety of other representations" (2017, 219). Kwisthout and Rooij (2013) point out that hierarchical predictive coding appears to be a matter of brain organization, while Bayesian inferences are a function or performance of that very brain. Fink and Zednik (2017) see Bayesian inference and hierarchical predictive coding as two methodologically distinct frameworks, the former of which is rooted in cognitive psychology, where it successfully captures human behavior, while the latter is used in theoretical neuroscience and is mostly concerned with information processing in the brain. Obviously, the relationship between the models is a problem that this paper cannot solve, or even discuss at length and greater detail. Suffice it to say that we believe that a combined model ("Bayesian predictive coding") should be most interesting and useful for the kind of question discussed here, as it combines actual neural facts and processes with a plausible probabilistic model of human behavior.

Processing Narratives and the Aesthetics of the Unexpected

It seems more than plausible to assume that general cognition and cognizing are not simply about the reaction to stimuli from the outside world. Rather, it has been
suggested that cognition is forward-looking in time in the sense that certain states-of-affairs are more expected than others (see section 2 above). This can be tested and shown for conscious cognitive operations, e.g., by asking subjects to continue a certain string of objects, such as in figure 1.

Figure 1: String of objects

Subjects will usually assume that the next symbol will be the face emoji. Similar operations also seem to be at work in subconscious cognition, e.g. in intuitions and spontaneous perception. Based on previous experience, cognizers tend to predict the likelihood of particular states. In the famous "dots illusion" discussed by Ramachandran and Rogers-Ramachandran (2008) subjects are presented with pictures of dots with different shading, as in figure 2.

Figure 2: Dots illusion (adapted from Ramachandran and Rogers-Ramachandran [2008])

Most cognizers perceive the dots on the left-hand side as concave and those on the right-hand side as convex. However, both left and right are one and the same picture, only turned by 180 degrees. If you flip the page, the formerly convex dots will be concave, and vice versa. In our previous experience, we have come to (statistically) associate a shade in the upper half of a body with a concave shape, and light in the upper half with a convex shape. This is due to the fact that in real life, in a three-dimensional space, light often comes from one source above and it lights the upper half of a given convex body (say, a ball), whereas the body of a concave object (say, a shell) will lead to some shade in the upper, inner part. Given these embodied experiences, we will tend to assume (based on statistical likelihood) that the objects on the left-hand side will be three-dimensional and concave (e.g., seashells seen from the inside), whereas the objects on the right-hand side will be seen as balls or such. Note that this is not a conscious operation or rational thought. This kind of reasoning happens unconsciously and is in fact even hard to influence by conscious thought. Even when we know that these two pictures are one and the same, we will still see concave and convex shapes. And just as we would be very surprised to find that the next symbol in (1) is not the face emoji, but a dollar sign ($), we are surprised to find that the dots are the same.
Probabilities, predictions, and surprise in the context of narrative texts may be discussed from at least two different angles. From the perspective of literary and cultural studies, we may look at expectations and surprise from an aesthetic or poetic point of view, i.e. in the discourse and deliberations of artists and critics. At least the term and concept of ‘surprise’ seems to have played a central role in many poetics, beginning with Aristotle, who sees surprise as a key element in the *peripeteia* (Aristotle 1995, Ch. V). Daniel Defoe advertises *Robinson Crusoe* (1719) as a story of Crusoe's *Life and Strange Surprizing Adventures*, and "Make it new" was the well-known slogan of Ezra Pound's modernist essay, published in 1935. But not all aesthetic or poetic practices appreciated or even called for the new, the surprising. Tyler (2006), for example, discusses the familiar, the conventional as a key element in Anglo-Saxon poetry. Other examples appear to be hard to come by; it is the lack of praise for surprise that characterized them rather than the explicit praise of the unsurprising and conventional.

One of the most extensive studies of probabilistic thinking in narratives is Kukkonen (2014; in press). She couches her work in the model of Bayesian inferences and distinguishes between three orders of probability design employed by authors. In the first order, we are looking at the trajectory plot and any prediction errors related to this. These, Kukkonen explains, are usually made or enacted by characters in the text and by readers of the text alike. They are usually noticeable in that characters explicitly comment on their surprise or the unexpectedness of a certain twist, event, or fact. The second-order probability design of a narrative runs through the whole text and is subtler and difficult to notice. Here, stylistic or linguistic elements guide the embodied reader to the precision of the predictions, and prediction errors. The third-level probability design of a text is more abstract and embeds the text in intertextual and generic dimensions. Here, the reliability of certain predictions takes center stage. There is also a cascade of feedback loops between predictions and prediction errors, so that lower-level predictions and prediction errors may feed back into higher-level predictions, which may have to be adjusted accordingly or which can be strengthened. At the same time, any prediction errors at higher levels also mean re-adjustments of lower-level predictions. Kukkonen offers a most comprehensive picture of the role of prediction and probabilities in narratives. Not only does she include contextual features, such as intertextuality, she also includes the reader and the characters of a text as cognizers making predictions and reacting to prediction errors. Ultimately, however, this is only part of probability designs created by authors. These designs are anything but random – they are complex and well-crafted systems that motivate and encourage particular predictions and that require prediction errors as motors that drive the narrative forward.

There is a second perspective on probabilities, predictions, and surprise that is worth mentioning: the perspective of cognitive science and artificial intelligence. Are there any general and possibly universal functions of expectation and surprise in cognition, in particular with regard to aesthetic pleasure? Why do these phenomena hold such a central position for many of us? Schmidhuber (2009), for example, claims that truly surprising (as well as interesting or beautiful) information is not just any new kind of information (e.g., white noise). It is neither fully predictable nor completely arbitrary. For Schmidhuber's theory of the aesthetic, surprising information must be new and
unexpected, but subjectively compressible. Schmidhuber also claims that "[s]ince short and simple explanations of the past usually reflect some repetitive regularity that helps to predict the future as well, every intelligent system interested in achieving future goals should be motivated to compress the history of raw sensory input in response to its actions, simply to improve its ability to plan ahead" (2009, 49). In other words, other things being equal, we would expect human agents to have an internal rewarding system for successful compression – to strive for the integration of formerly unknown regularities in the outside world, and to appreciate new and surprising, but compressible information as it sharpens our cognition. This is not to be confused with external rewards, such as the pleasant feeling one might get upon hearing a certain song or reading a particular poem. These external rewards may be explained by social psychology, for example, and they also play some role in our evaluation of beauty and interestness. In contrast, Schmidhuber maintains, the internal reward for compressing new and surprising information into regular and predictable patterns is what lies at the heart of 'actual' beauty and aesthetics. This appears to be perfectly compatible with what Meir Sternberg identifies as the universals of narrative: "suspense, curiosity, and surprise, each encoding a distinct functional operation of the mind within narratives' overall inter-sequencing, i.e. the dynamics of prospection, retrospection, and recognition, respectively" (2003, 327).

A Case Study in Processing Idiom Variability

As we have mentioned at the beginning, empirically investigating probability design, predictive coding and surprise in actual longer texts prove to be extremely difficult. For this reason, we decided on a kind of sandbox or miniature study that exemplifies and tests the general principles outlined above, and that can be extended and projected onto longer and more complex texts. This sandbox study investigates the variability and use of a set of idiomatic expressions in English, such as "He is not the sharpest tool in the shed." Idiomatic expressions provide us with excellent data for investigating how recipients process and interpret creative, i.e. surprising and unexpected language use. Idioms are often based on schematic constructions that contain open slots which can be filled in by the speaker. Thus, they allow for a great deal of variation which needs to be decoded and processed by the hearer (cf. Hilpert 2014, 6). Our case study aims at investigating whether and how recipients interpret novel (surprising) variants of an idiomatic construction and to what extent the open slots can be filled and interpreted until the reader deems the idiomatic expression incomprehensible.

Based on the idiomatic expression X is not the sharpest tool in the shed, we first conducted a corpus search in iWeb and NOW\(^1\) and identified two alternations of this idiom that already are in frequent use, i.e., [X] is not the sharpest knife in the drawer and [X] is not the brightest bulb in the [Q]. The frequencies (per million words) are shown in figure 3. For the latter example, the [Q] slot is most frequently filled with box

---

1 The iWeb corpus contains 14 billion words and the NOW (News on the Web) corpus contains 7.1 billion words. Both corpora are web-based (https://corpus.byu.edu/corpora.asp).
(in iWeb ca. 26.09%; in NOW ca. 18.18%) (see figure 4). All other options for the Q slot did not occur as frequently.

Figure 3: Frequencies (per million words) for three idiom varieties in iWeb and NOW

Figure 4: Percentage of slot option in the box compared to all other options Q

These three variants of the idiomatic expression are all derived from the underlying construction [NP BE NOT the sharpest/brightest N PrepP], which exhibits several empty and/or variable slots, and all variants share the same meaning, i.e., the subject is not particularly intelligent. An obligatory variable element of the underlying construction is the verb to be (present, past, third-person etc.). The negator not is invariable, while the nouns and the adjective can be altered.

The inflectional form of to be does not have to be limited to the present tense form, occurrences in the past tense are licensed by the construction as well. However, other
verbs cannot be used since their use would alter the meaning of the idiom. The negator cannot be left out for the same reason (see examples [1] and [2]).

(1) "X does not have the sharpest tool in the shed.
(2) "X is the sharpest tool in the shed.

Altering only the nouns still yields the meaning of the idiom as being 'not particularly intelligent.'

(3) X is not the sharpest tack in the box.
(4) X is not the brightest star in the sky.
(5) X is not the sharpest arrow in the quiver.

A more abstract notion of the idiom can be reached, and thus a more abstract underlying schema if we change the whole noun phrase. In (4) we have already altered the adjective, bright vs. sharp, but we can also use adjectives that are not associated with intelligence. However, in those cases we also need to change the nouns accordingly since the adjective needs to fit the noun in the literal sense. Examples for such altered NPs are given in (6) and (7).

(6) X is not the sweetest candy in the box.
(7) X is not the hottest marshmallow in the fire.

Considering these other NPs, we can derive a more general meaning of the idiom as 'X is not particularly Y.' Thus, we can describe the underlying schematic construction as [NP BE NOT the ADJ\text{superlative} N Prep the N\text{sg}]. This will also be used as the template for testing the processing of this idiom and its variants. In order to develop a reliable basis for the test items, several conditions have been determined for the individual slots of the construction:

<table>
<thead>
<tr>
<th>Slot</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>• animate referent</td>
</tr>
<tr>
<td></td>
<td>o most frequently human</td>
</tr>
<tr>
<td></td>
<td>o less frequently animal</td>
</tr>
<tr>
<td>BE NOT the</td>
<td>• none</td>
</tr>
<tr>
<td>ADJ\text{superlative}</td>
<td>• attributive use for animate referent possible</td>
</tr>
</tbody>
</table>

2 The non-negated forms are neither attested for in the corpora nor by interviewed native speakers. Even though we are aware that we are lacking positive evidence for the impossibility of this construction, we believe it is highly unlikely to occur.

3 "The sweetest tool in the shed" is not meaningful, since "tools" cannot be "sweet." See also slot conditions below.
Table 1: Slot conditions

| N Prep the N_sg | • has to fit the [ADJ superlative] in the literal sense  
| | • has to be internally coherent and logical |

The initial NP slot is already quite variable as long as the condition of the animate referent is met and filling it with non-animate referents would not lead to insightful results concerning the recipients' processing. Further alternations in the last slot, [N Prep the N_sg], would not be productive either since this part appears to be semantically irrelevant and syntactically non-obligatory, i.e. it can be elided. As a matter of fact, this part only decreases the opacity of the idiom (Taylor 2012, 73), as can be seen in examples (8) and (9), which are both parsable without the [N Prep the N_sg], but where this part facilitates unambiguous processing.

(8) X is not the sharpest tool in the shed.
(9) X is not the brightest bulb in the box.

In order to examine how cognizers react to unexpectedness and how they deal with hitherto unknown data, we presented subjects with several 'new,' made-up variations of this idiom. We chose to vary the ADJ superlative slot as the central element and created some test items in which the condition of applicability to animate referents is actively disregarded. By using adjectives that are not conventionally used as attributes for animate referents, these items should be deemed incomprehensible by the recipients. An example of such a test item is given in (10).

(10) *X is not the roundest ball in the basket.

Round is not conventionally used as an attribute to describe animate referents, thus the slot condition of ADJ superlative is not fulfilled. Even though the other slot conditions are met, example (10) should not be interpretable. This systematic choice of the slot that is to be varied should thus show to what extent speakers/hearers will be able to use their constructional knowledge of the underlying schema as a predictive force in order to encode and decode the idiom.

Furthermore, we incorporated several items that are not merely varied by slot filling, but that also contain metaphor blends. Since the metaphors used in our study all include an adjective themselves, these items test whether speakers/hearers will focus on the adjective itself or whether they only interpret it in connection with the inserted metaphorical concept. An example of such an item is given in (11).

(11) X is not the biggest fish in the sea.

In total, in our study there are three different groups with various subgroups:
The study was implemented via LimeSurvey, provided by Osnabrück University. The access link to the online questionnaire was shared with several universities in the United States, Great Britain, Australia, the LinguistList platform, and individual native speakers, all of which were encouraged to share the link with other native speakers. The study was made publicly accessible, i.e., participants were not selected individually beforehand. Since the study is primarily aimed at English native speakers, the initial question provided the following choices to enable categorization according to English speaking experience:

- native speaker
- more than 8 years speaking experience
- 6-8 years speaking experience
- 3-5 years speaking experience
- 0-2 years speaking experience

The structure of the questions stays the same throughout the test. For every item, the participants are presented with a variant and the question "Does this expression make sense to you?" with 'Yes' and 'No' as the only answers. If a participant chooses 'No,' they are redirected to the next variant. If they choose 'Yes,' however, the study redirects them to a subquestion: "What could this expression mean? Please paraphrase it and give a brief

---

4 181 participants registered as native speakers, nine as 'more than 8 years' and two as '6-8 years.' As their answers were in no way different from that of the native speaker group, the results for all subjects (192) were lumped together.
explanation for your interpretation." The participants are then given two text fields, one of which is labeled "This expression could mean:" and the other "Explanation:". The first text field thus asks them for their interpretation, e.g., by paraphrasing the variant, while the second one gives them the opportunity to explain their interpretation. Both text fields are vital for the interpretation of the study. However, filling in these text fields is not required in order to proceed with the questionnaire.5

In total, 192 speakers participated in this study. However, not all results proved to be useful: out of all participants, 104 completed the questionnaire; 88 submitted only partial answers. Participants who only answered the first question about their proficiency level or the first and second question were excluded in the analysis and interpretation of the study results. This was true for 78 participants. Of the ten remaining participants, we did not exclude those who did not fully complete the test, since even single answers to the first few items of the study appeared to be useful. As we did not assume (or test for) quantitative relations between the individual test items, the test works with any number of completed items. In the following, the group of participants who provided useful test results will be referred to as A-group.

- Total number of participants: 192
- Excluded participants: 78
- A-group: 114 (native speakers: 99; complete tests: 104)

Because the A-group still provided a huge amount of data, especially considering the free text fields, it would go beyond the scope of this paper to discuss every answer individually. Instead, particularly interesting results will be discussed further, as well as observations relevant to the key question of this paper.

Nevertheless, a brief overview of the results and different interpretations will be provided in the following. Of course, a summary of answers is necessarily an interpretation to a certain degree. On a few occasions, participants gave interpretations without negation ('the strongest' instead of 'not the strongest'), which were translated into the negated version of the interpretation unless the participant specifically emphasized that the ADJ\textit{superlative} still applied to X, despite the negation in the variant. Furthermore, it should be mentioned that if one participant gave two interpretations for one variant, both are counted and listed individually, which is why the total number of interpretations may exceed the number of 'Yes'-answers. Since the following overview represents all participants, even those who did not complete the questionnaire, the number of answers, in general, may vary between 114 and 104.

---

5 All test items and an example of the setup are given in Appendix A.
<table>
<thead>
<tr>
<th>Item</th>
<th>(Yes/No) Answers</th>
<th>Interpretations summarized</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. She is not the sharpest tool in the shed.</td>
<td>(113/1) 99,1% Yes</td>
<td>• Not intelligent (109)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not the best (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not useful (1)</td>
</tr>
<tr>
<td>2. She is not the sharpest knife in the drawer.</td>
<td>(100/14) 87,7 % Yes</td>
<td>• Not intelligent (90)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not very effective (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not quick (1)</td>
</tr>
<tr>
<td>10. She is not the brightest crayon in the box.</td>
<td>(75/29) 72,1 % Yes</td>
<td>• Not intelligent (54)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Boring, does not stand out, not noticeable/interesting/vibrant (16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not socially engaging/convivial (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not optimistic (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not creative (1)</td>
</tr>
<tr>
<td>13. She is not the brightest bulb in the box.</td>
<td>(78/26) 75 % Yes</td>
<td>• Not intelligent (75)</td>
</tr>
<tr>
<td>3. She is not the quickest bunny in the field.</td>
<td>(67/44) 60,4 % Yes</td>
<td>• Not intelligent/mentally perceptive (44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not fast (physically/running speed) (14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not the best/not fit for the task (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Lacking social abilities (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does not jump to conclusions (1)</td>
</tr>
<tr>
<td>12. He is not the straightest line in the drawing.</td>
<td>(35/69) 33,7 % Yes</td>
<td>• Gay (13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Crooked/immoral/dishonest (11)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Weird/Crazy/Odd/Messed up (9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Unconventional/Walks a different path (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not accurate (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not smart (1)</td>
</tr>
<tr>
<td>20. He is not the coolest drink in the fridge.</td>
<td>(28/76) 26,9 % Yes</td>
<td>• Boring/lame/nerdy (16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not attractive (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not calm, short-tempered (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not the best (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not sophisticated/poised (1)</td>
</tr>
</tbody>
</table>
### Table 3.1: R-Regular

<table>
<thead>
<tr>
<th>Item</th>
<th>.rl</th>
<th>Interpretations summarized</th>
</tr>
</thead>
</table>
| 21. He is not the hottest steak on the grill.                        | 26/78 | • Not attractive (sexually) (21)  
|                                                                      | 25 % Yes | • Not interesting/important (relevance) (2)                                                    |
| 6. He is not the prettiest flower in the field.                     | 81/29 | • Ugly/not particularly pretty (74)  
|                                                                      | 73,6 % Yes | • Not desirable (while not referring to physical appearance) (1)  
|                                                                      |          | • Does not stand out (3)                                                                    |
| 15. He is not the strongest boxer in the ring.                      | 63/41 | • Physically weak, not strong, defenseless (31)  
|                                                                      | 60,6 % Yes | • Not the best (competitor/candidate), not skilled, lacking qualities, inadequate, not influential/powerful, emotionally weak, loser (11)  
|                                                                      |          | • At a disadvantage (physically or ability-wise), about to lose (4)  
|                                                                      |          | • Not smart (1)                                                                            |
| 26. He is not the calmest pond in the park.                         | 36/68 | • Easily agitated/angered/upset, short-tempered, not calm (30)  
|                                                                      | 34,6 % Yes | • Not calm: Crazy/wild/energetic, loud or anxious/giddy (9)                                    |
| 18. He is not the bravest lion of the pack.                         | 69/35 | • Not brave, scared/wimpy (62)  
|                                                                      | 66,3 % Yes | • Not the alpha male, soft-spoken, will not step up to be a leader (5)                        |
| 8. He is not the whitest sheep of the flock.                        | 40/68 | • Not moral/pure/good/innocent, bad reputation, "track record of bad decisions" (33)  
|                                                                      | 37 % Yes | • Outsider:"Black sheep"/not normal (5)  
|                                                                      |          | • Not clean (1)                                                                            
<p>|                                                                      |          | • Not Caucasian (racist undertone) (1)                                                      |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 11. He is not the biggest fish in the sea. | (68/36) 65,4 % Yes | - Not important/powerful/successful, not an authority figure (51)
- Not tall/strong (7)
- Not the best catch (6, of which 1 referred 'big' to the size of male genitalia)
- Not the best (3)
- Not wealthy (2)
- Not smart/wise (2)
- Shy (1) |
| 16. She is not the sweetest pea in the garden. | (42/62) 40,4 % Yes | - Rude/unpleasant, not very nice/kind (37)
- Not the best (1)
- Not young anymore (1) |
| 22. She is not the lightest feather in the pillow. | (26/78) 25 % Yes | - Overweight (23)
- Not prim (1)
- Not the sweetest (1)
- Stiff (1) |
| **Table 3.2: M_Metaphor blend** | **MM_Metaphor blend: malaphor** | | |
| 9. He is not the strongest scent in the air. | (10/97) 9,3 % Yes | - Subtle, not very noticeable (6)
- Not influential/"alpha male" (2)
- Not smelly (1)
- Smelly (1)
- Not smart (1) |
| 25. He is not the highest horse in the stable. | (16/88) 15,4 % Yes | - Not tall (5)
- Not (the most) arrogant (4)
- Not important (3)
- Not talented, "idiot" (2)
- Not confident (1)
- Low standards/morals (1) |
<table>
<thead>
<tr>
<th>Item</th>
<th>(Yes/No) Answers</th>
<th>Interpretations summarized</th>
</tr>
</thead>
</table>
| 4. She is not the fluffiest pillow on the couch. | (28/82) 25.5% Yes | - Not kind/nice/charming (8)  
- Thin (4)  
- Tough, not soft/cuddly (3)  
- Not the comfiest (2)  
- Not the most comfortable to be with (2)  
- Not competent/good at something (2)  
- Erotic undertone (1)  
- Not good (1)  
- Not the prettiest (1)  
- Not smart (1) |
| 5. He is not the whitest cloud in the sky. | (26/84) 21.8% Yes | - Ethically/morally questionable person, guilty (5), Not innocent/pure (7) (12 total)  
- Not in a good mood/gloomy/depressed (9)  
- Not clean (2)  
- Racist undertone (1) |
| 7. She is not the greenest leaf on the tree. | (36/74) 32.7% Yes | - Old, not young and vibrant (19)  
- Experienced (9)  
- Not very capable/good at sth. (2)  
- Room to grow intellectually/spiritually (1)  
- Not pretty (1)  
- Not special (1)  
- Tired/Worn down (1)  
- Not smart (1)  
- Not fruitful (1)  
- Not natural/conservation-minded (1) |
| 14. He is not the ripest apple in the basket. | (28/76) 26.9% Yes | - Too old or too young and naïve, not mature, not in their prime (11)  
- Not ready (4)  
- Ugly, not attractive (3)  
- Not smart (3)  
- Corrupt/tarnished/mean (3)  
- Not the best (choice) (2)  
- Not put together (1) |
| 17. She is not the roundest ball in the basket. | (11/93) 10,6 % Yes | • Not flawless/functional (3)  
• Different/unconventional (2)  
• Skinny, "beanpole" (2)  
• Not a well-rounded human being (in terms of skills, interests, etc.) (1) |
| 19. She is not the glassiest window in the house. | (5/99) 4,8 % Yes | • Not attractive/stylish (2)  
• Not clean (2)  
• Not transparent (1) |
| 23. He is not the widest angle in the graph. | (6/98) 5,8 % Yes | • Skinny or narrow-shouldered (2)  
• Not smart (2)  
• Limited in outlook (1)  
• Not the most open-minded (1) |
| 24. She is not the briefest letter in the mail. | (22/82) 21,2 % Yes | • Talks a lot (21)  
• Has a lot of backstory/personal experience (1) |
| 27. She is not the crunchiest cookie in the box. | (8/96) 7,7 % Yes | • Tough, not soft (2)  
• Less than ideal/Not the best (1)  
• Not pleasant to be around (1)  
• Not smart (1) |
| 28. She is not the hardest brick in the wall. | (26/78) 25 % Yes | • Not very resistant, not strong in character, weak/easily hurt (12)  
• "Softie," not hard-hearted, forgiving (3)  
• Not obstinate/stubborn, open-minded, lenient, open to compromise/bribe (6) |

Table 3.3: T Test group

The explanations given by the participants for their interpretations are not included in table 3.1-3.3 due to their extensive variation. In the following, explanations will only be included in cases where they seem useful.

To allow for a better comparison between the three groups listed above, the average number of 'Yes'-answers for each group was calculated and is given in percentage in table 3.4.
### Table 3.4: Average number of 'yes' for each group

<table>
<thead>
<tr>
<th>Group/Subgroup</th>
<th>Average Percentage of 'Yes'-answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_Regular</td>
<td>59,59</td>
</tr>
<tr>
<td>RPO_Regular polysemes original</td>
<td>83,48</td>
</tr>
<tr>
<td>RP_Regular polysemes</td>
<td>36,50</td>
</tr>
<tr>
<td>RL_Regular literal</td>
<td>58,78</td>
</tr>
<tr>
<td>M_Metaphor blend</td>
<td>27,15</td>
</tr>
<tr>
<td>M_Metaphor blend: regular</td>
<td>41,95</td>
</tr>
<tr>
<td>M_Metaphor blend: malaphor</td>
<td>12,35</td>
</tr>
<tr>
<td>T_Test group</td>
<td>18,20</td>
</tr>
</tbody>
</table>

**Interpretation**

In this section, we will offer an interpretation of the results of our empirical study, presented in section 4. As expected, our subjects had only few problems with the RPO_group, i.e. the handful of expressions that are fairly common and well-known to native speakers. Unexpected were only the results of RPO_10. *(brightest crayon)*. Even though the percentage of 'Yes'-answers (72,1%) is close to the one of RPO_13., sixteen out of 75 interpretations showed some variation of 'X is boring, does not stand out, not noticeable/interesting/vibrant.' Due to the lower corpus frequency of this variant compared to the other RPO_variants (cf. section 4), it is possible that participants never heard this variant before. Due to RPO_10. preceding RPO_13. in the test, participants who have never heard the *brightest crayon*-variant were not reminded of the polysemous character of 'bright' by the more common variant with *brightest bulb*. When the synonym of 'smart' is not triggered with this adjective, this item necessarily becomes a T_variant, which would then justify a different interpretation method, which can be found in the T_group as well: a striking feature in many T_group results surfaces when the participants' explanations are taken into account: Even though most of these items show massively scattered interpretations (cf. table 3.3, *fluffiest pillow*), the participants often seemed to use the same approach to ascribe meaning to the variant. In many explanations, the first N-slot was mentioned alongside the ADJ: 

1. "a ripe apple reaches its finest 'flavor'/sweetness." (id: 531, T_14.)
2. "balls by definition are round. A ball that is not round is imperfect." (id: 180, T_17.)

---

6 Frequencies (per million words): 0.0019 (iWeb); 0.0005 (NOW).
Nevertheless, these observations can also be made for variants of the R_group, where participants included the N-slot in their explanation. However, while the N-slot was mentioned on various occasions in R_variant explanations, it did not seem to heavily influence the interpretation in most cases. For item 26 (calmest pond), for example, the word 'pond' (or 'water') was mentioned in 14 out of 31 given explanations. However, the participants only provided two kinds of interpretations, either that X is 'easily agitated/angered/upset, short-tempered, not calm' (30x) or 'crazy/wild/energetic, loud or anxious/giddy' (9x), which do not build upon a possibly metaphorical concept of a pond, but are rather based on the negated version of 'calm.' Participants thus did not necessarily understand 'pond' as a metaphor but rather claimed that X is not calm as ponds are. Similar to ' sharpest knife,' the N-slot functioned as a filler word with little semantic relevance. The constructional knowledge of the participants thus outweighed possible metaphorical interpretations. The T_variants, however, showed interpretations with a significant tendency to not only focus on the adjective at hand but building up a complete metaphorical concept as in (1) or (2).

Considering that most 'Yes'-answers for T_variants followed this approach, it can be concluded that participants who proved to be more progressive in terms of accepting new constructions followed the same schema for interpretation. The new and broader meaning side, which most of them used, would be 'X = not very ADJ N' with the condition that [ADJ N] works as a metaphorical concept which can be applied to humans. This would also explain the numerous different interpretations given for the T_variants: When [ADJ N] does not present a conventional metaphor, the participants had to come up with their own interpretation of the new metaphor and ascribe meaning to the idea of 'X is a fluffy pillow.' Since context was neither provided nor known for these metaphors, the results had to be as scattered as they are. Our previous assumption that these results require more processing than the R_group is underlined by the low results the T_variants scored. 45 out of the 104 participants who completed the whole questionnaire did not answer a single variant out of the T_group with 'Yes.'

Both the T_group (with non-sensical expressions) and the MM_group proved to be worst in terms of acceptability with 18% and 12%, respectively. The MR_group, in comparison, scored rather high results with almost 42%. The highly conventionalized metaphors (adjective + noun) used in these items were mostly recognized and used in the interpretations, which is why it can be concluded that the approach chosen by most participants for T_variants was also used for the MR_group. Moreover, this approach seems to be more effective for the MR_items than for the T_items. When the metaphorical concept is known and recognized, 'X is not ADJ N' can be applied easily, without additional information needed. Especially when comparing items T_5. and MR_8., it becomes clear how conventional knowledge influenced the participants and their interpretations: both items had the ADJ superlative-slot filled with the adjective 'white,' which was deemed a breach of the ADJ superlative Condition. Even though it can be used as a reference to skin color, this was only mentioned once in the interpretation of each item (id: 33 cloud, id: 81 sheep). Taking into consideration only the conditions established previously, the number of 'Yes'-answers should be the same for both variants, since the ADJ superlative Condition was neglected the same way in each. However, while the MR_variant ( whitest sheep) scored 37 % 'Yes'-answers, the T_variant
(whitest cloud) only scored 21.8% 'Yes'-answers. The highest-scoring interpretations for the T_variant were for white = 'ethically/morally questionable person, guilty' (5x) and 'not innocent/pure (7x), and for white cloud = 'not in a good mood/gloomy/depressed' (9x) (based on participants’ explanations). In this specific example, it should be considered that the adjective 'white' in itself has a highly conventional, symbolic connotation, of which the 'black sheep' metaphor may have been derived. This would explain the high results for this interpretation in particular. The interpretations featuring the 'white cloud' imagery are far more promising, however, considering that they followed the previously mentioned meaning side 'X is not ADJ N,' as the following explanations can attest:

1. "Clouds that are white aren't storm clouds. Storms are often associated with darkness and depression." (id: 63)
3. "Grey clouds = gloomy." (id: 490)

Those participants who chose this way of interpretation agreed upon one meaning, which is why 'white cloud,' for example, seems to be a better metaphorical concept than the more ambiguous 'fluffy pillow.' Nevertheless, it still scored lower results than the MR_variant, since 'white cloud' is not a conventional metaphor. For both variants, however, the prototypical interpretation 'X is not ADJ superlative' did not work, which is why participants reverted to using 'X is not ADJ N' for their interpretation. This method requires conceptual integration (see Fauconnier and Turner 2002), which may not have been possible or acceptable for all participants, due to the new and – more importantly – unexpected processing load. The low results in the MM_group, moreover, prove that this interpretation method has its limits as well and can only be applied if the metaphor is conventionally used to describe humans as well.

The RL_items (where the adjective may actually be used literally for an animate individual) were more successful: they proved to be the second-strongest subgroup in the test, following only the RPO_variants. These results were expected and are in line with our hypothesis. The low results yielded by the RP_group are interesting, however, considering that their polysemous character is closer to the variants in RPO_ than the ones in the RL_group. RP_3. (quickest bunny) still shows a lot of 'Yes'-answers with 60.4%, while the other three items lie between 25-33.7%. The fact that these variants did not score a higher number of 'Yes'-answers might be based on the same reason why the idiom with the polysemic condition has not yet evolved into a full snow-clone: the spontaneous production and, as our results have shown, also the interpretation of the adjective's polysemous character poses a mental workload not processable in a speech situation. When the polysemic character is not recognized, RP_12., 20. and 21. become, similar to RPO_10., T_variants and have to be interpreted with 'X is not ADJ N.'

This first empirical test shows that cognizers are prepared to deal with new and surprising input and that they are willing to invest some effort into the processing of unexpected data. Here, in this case, hearers may first expect the regular common idiom. If this is not available, they appear to expect other patterns, albeit to varying degrees, starting with regular literal expressions through conventionalized metaphors to new and
unknown metaphorical concepts. Since metaphors are a common device in literature to evoke certain mental images (cf. Croft and Cruse 2004, 195), speakers are familiar with conventionalized metaphors. In addition to commonly used mental images, metaphors are also used to ascribe meaning to unfamiliar expressions. In all of these situations, numerous participants did not reject the input as nonsensical, but rather applied universal cognitive mechanisms in an effort after meaning principle (see Bartlett 1932). The more common the expression they encounter (the more frequent the polyseme or blend), the more likely they are to find the structure acceptable and to offer an interpretation. But this also seems to have its limits: both the Test Group (T_) and the malaphors (MM_) show that a significant number of hearers are not willing (or capable) to apply these principles to all sorts of input. When a certain threshold is reached and input becomes incompressible (or only compressible with significant effort) hearers are more likely to reject the utterance.

Conclusion

In this brief paper, we have attempted to show that a model of predictive Bayesian coding may not only apply to basic and intuitive cognition (such as visual perception) but that it may also be interesting for more or less complex linguistic material. Speakers/hearers appear to process linguistic information in a forward-looking, predictive way, and their perception of linguistic input is guided by these statistical inferences. When the input is unexpected to some degree and reaches a particular threshold, hearers utilize general cognitive principles in order to arrive at a plausible interpretation, which in turn updates their predictive system for new input. In other words, speakers who encounter 'new' but compressible idioms find it increasingly easier to process the new input. But this only applies to a certain degree – once a particular point is reached and input is no longer parsable (compressible), hearers reject the structures as nonsensical.

As we have pointed out at the beginning, these processes are difficult to test for more complex linguistic input such as narrative texts. Yet, we believe that even longer texts and narratives are ultimately processed in similar ways. Readers enjoy the processing of new and surprising information as it provides them with stronger cognitive processing skills and higher predictive accuracy. At the same time, they will probably also either reject input which is incompressible or will revert to more explicit and conscious processing strategies that involve more facts and aspects outside the text itself, such as cultural or social background or aesthetic/poetic principles and practices. Needless to say, here we are facing the infamous (and yet unresolved!) micro-macro problem when applying cognitive or linguistic ideas, mechanisms or principles to literary works. Cognitive science, neuroscience, and most of linguistics are usually concerned with the micro-level, i.e. they concentrate on 'smaller' and more clearly delineated phenomena such as word recognition, color vision, or morphemes. This has to do with both methodological rigor and with particular ways of 'doing science,' i.e. an orientation towards deontic-nomological modes of explanation. Mainstream, structural linguistics usually stops at the sentence level, for example. This is not to say that there are no structural units beyond the sentence.
The 1970s and 1980s saw a massive interest in what was termed text linguistics (Dressler 1978; Beaugrande and Dressler 1981; Werlich 1976; Kallmeyer 1977, to name but a few). The idea behind this was and is that even larger units of language, such as paragraphs and texts can be described in structural terms with the same methodological rigor and scientific interest as traditional linguistic structures such as phonemes, morphemes, and sentences. Unfortunately, it appears as if this branch has always led a life in the shadows, at least to some extent. The current paper, with its case study based on a pervasive micro-phenomenon, is a plea for renewing our linguistic and cognitive interest in the larger units such as paragraphs and texts. We need to further explore what is called the micro-macro line in order to be able to discuss in what ways phenomena that are observable on the micro level also apply to the macro level of language. As the null hypothesis we would suggest that the macro-level should not be fundamentally different from what we observe on a smaller scale. But in order to show this, we will have to develop new tools and methods that allow us to study, for example, concrete reader behavior in relation to larger (literary) texts. Eye-tracking studies (Vančová 2014), empirical studies of literature (e.g. Miall 2006) and Kukkonen's predictive coding approach (in press) are certainly steps in the right direction and we should further bridge the gap between micro and macro. We hope this paper offers at least some contribution towards that.

Future research will have to probe into our hypotheses by developing test batteries for larger chunks of texts and narratives.

Works Cited


Appendix A
List of All Items of the Study

1. She is not the sharpest tool in the shed.
2. He is not the sharpest knife in the drawer.
3. She is not the quickest bunny in the field.
4. She is not the fluffiest pillow on the couch.
5. He is not the whitest cloud in the sky.
6. He is not the prettiest flower in the field.
7. She is not the greenest leaf on the tree.
8. He is not the whitest sheep of the flock.
9. He is not the strongest scent in the air.
10. She is not the brightest crayon in the box.
11. He is not the biggest fish in the sea.
12. He is not the straightest line in the drawing.
13. She is not the brightest bulb in the box.
14. He is not the ripest apple in the basket.
15. He is not the strongest boxer in the ring.
16. She is not the sweetest pea in the garden.
17. She is not the roundest ball in the basket.
18. He is not the bravest lion of the pack.
19. She is not the glassiest window in the house.
20. He is not the coolest drink in the fridge.
21. He is not the hottest steak on the grill.
22. She is not the lightest feather in the pillow.
23. He is not the widest angle in the graph.
24. She is not the briefest letter in the mail.
25. He is not the highest horse in the stable.
26. He is not the calmest pond in the park.
27. She is not the crunchiest cookie in the box.
28. She is not the hardest brick in the wall.
Appendix B

Example Question Design

She is not the fluffiest pillow on the couch.

Does this expression make sense to you, i.e. would you understand it in a conversation?

Yes  No

She is not the fluffiest pillow on the couch.

What could this expression mean? Please paraphrase it and give a brief explanation for your interpretation.

This expression could mean: _________________________

Explanation: _________________________

REMINDER: There are no right or wrong answers. I just want to know how you would understand this expression.