COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

CATEGORIZATION, COMPARISON AND CONVENTIONALIZATION IN METAPHOR COMPREHENSION

DIPLOMA THESIS

2015

Bc. Zdenko Kohút

COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

CATEGORIZATION, COMPARISON AND CONVENTIONALIZATION IN METAPHOR COMPREHENSION

DIPLOMA THESIS

Study programme:	Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study:	2503 Cognitive Science
Training work place:	Department of Applied Informatics
Supervisor:	RNDr. Barbora Cimrová, PhD.

Bratislava 2015

Bc. Zdenko Kohút

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Bc. Zdenko Kohút
Study programme:	Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study:	9.2.11. Cognitive Science
Type of Thesis:	Diploma Thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Categorization, comparison and conventionalization in metaphor comprehension

Aim:
1. Evaluate two different theoretical views on metaphor comprehension – the categorization and the comparison view – with regard to the process of conventionalization.
2. Consider their consistency with the relevance theory.
3. Test the theories using an empirical method.

Literature:
Bowdle & Gentner 2005. The Career of Metaphor. Psychological Review 112, 193-216.
Gentner D. 1983. Structure-Mapping: A Theoretical Framework for Analogy. Cognitive Science 7, 155-170.
Glucksberg & Haught 2006. On the Relation Between Metaphor and Simile: When Comparison Fails. Mind & Language 21, 360-378.

Annotation: Metaphor is almost omnipresent in natural language and thought. There are two influential competing views on metaphor comprehension. First one considers it a categorization. The other treats novel metaphor comprehension as a comparison but conventional metaphor comprehension as a categorization. This is usually tested using nominal metaphors and similes as tropes inducing figurative categorization and comparison, respectively.

Keywords: metaphor, simile, categorization, conventionalization, relevance

Supervisor:	RNDr. Barbora Cimrová, PhD.
Department:	FMFI.KAI - Department of Applied Informatics
Head of department:	prof. Ing. Igor Farkaš, PhD.
Assigned:	30.08.2014
Approved:	28.04.2015

prof. Ing. Igor Farkaš, PhD.
Guarantor of Study Programme

Student

Supervisor

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Zdenko Kohút
Študijný program: kognitívna veda (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.2.11. kognitívna veda
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Categorization, comparison and conventionalization in metaphor comprehension
Kategorizácia, prirovnávanie a konvencionalizácia v chápaní metafor

Cieľ:
1. Prehodnoťte dva rozličné teoretické pohľady na porozumenie metaforám – kategorizačný a porovnávací prístup – s ohľadom na proces konvencionalizácie.
2. Zvážte konzistenciu týchto teórií s teóriou relevancie.
3. Testujte tieto teórie empiricky.

Literatúra:
Bowdle & Gentner 2005. The Career of Metaphor. Psychological Review 112, 193-216.
Gentner D. 1983. Structure-Mapping: A Theoretical Framework for Analogy. Cognitive Science 7, 155-170.
Glucksberg & Haught 2006. On the Relation Between Metaphor and Simile: When Comparison Fails. Mind & Language 21, 360-378.

Anotácia:
Metafora je takmer všadeprítomnáv prirodzenom jazyku a myslení. Existujú dva významné súperiace pohľady na porozumenie metaforám. Jeden považuje metaforu za kategorizáciu. Druhý nazerá na chápanie nových metafor ako na prirovnávanie, ale na porozumenie konvenčným metaforám ako na kategorizáciu. Obvyklou empirickou paradigmou je použitie nominálnych metafor a prirovnaní ako trópov indukujúcich figuratívnu kategorizáciu resp. prirovnávanie.

Kľúčové slová: metafora, prirovnanie, kategorizácia, konvencionalizácia, relevancia

Vedúci: RNDr. Barbora Cimrová, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: prof. Ing. Igor Farkaš, PhD.

Dátum zadania: 30.08.2014

Dátum schválenia: 28.04.2015

prof. Ing. Igor Farkaš, PhD.
garant študijného programu

.....
študent

.....
vedúci práce

Acknowledgment

I would like to thank to my supervisor Barbora Cimrová for friendly approach and for providing me with participants for the study.

Abstract

There are two strong competing views on the nature of metaphor. First, represented by the class-inclusion theory proposes that metaphors are processed as categorizations. The other view, represented by the career of metaphor theory, claims that novel metaphors are processed as comparisons and only conventional metaphors are processed as categorizations. Based mainly on the career of metaphor theory and the relevance theory, we were able to describe the difference between categorization and comparison in metaphor comprehension in simple terms of number of properties shared by the concepts. We proposed that as the difference between the number of properties encoded by the abstract metaphorical and the target concept increases, the more likely will a metaphor be considered a categorization. Reversely, as the number of properties encoded by the abstract metaphorical concept approaches the number of properties encoded by the target, the more likely will a metaphor be considered a comparison.

Hypotheses implied by this proposal were tested experimentally. Our results are consistent with our predictions concerning novel metaphors. Unfortunately, the variance in the data did not allow us to draw any conclusions about conventional metaphors.

Additionally, we were concerned with neural aspects of figurative language processing in typical and non-typical populations and suggested a simple way to describe figurative language processing in schizophrenia and autism spectrum disorders consistent with our proposal.

Keywords: metaphor, simile, categorization, conventionalization, relevance

Abstrakt

V súčasnosti rozoznávame dva silné, súperiace pohľady na metaforu. Prvý, reprezentovaný takzvanou class-inclusion teóriou tvrdí, že metafory sú spracovávané ako kategorizácie. Druhý pohľad, reprezentovaný career of metaphor teóriou tvrdí, že nové metafory sú spracovanávané ako prirovnania a iba konvenčné metafory sú spracovávané ako kategorizácie. Predovšetkým pomocou career of metaphor teórie a teórie relevancie sme boli schopní popísať rozdiel medzi kategorizáciou a prirovnaním na základe množstva vlastností zdieľaných konceptmi. Navrhli sme, že čím väčší je rozdiel medzi množstvom vlastností kódovaných abstaktným metaforickým konceptom a target konceptom, tým je pravdepodobnejšie, že metafora bude chápaná ako kategorizácia. Naopak, čím sa množstvo vlastností kódovaných abstraktným metaforickým konceptom približuje množstvu vlastností kódovaných target konceptom, tým je pravdepodobnejšie, že metafora bude chápaná ako prirovnanie.

Hypotézy vyplívajúce z tohto návrhu sme testovali experimentálne. Naše výsledky sú konzistentné s našimi predikciami týkajúcimi sa nových metafor. Nanešťastie, variancia v dátach nám nedovolila potvrdiť či vyvrátiť naše predpoklady týkajúce sa konvenčných metafor.

V rámci práce sme sa taktiež zaoberali neurálnymi aspektmi spracovania obrazného jazyka u typickej a netypickej populácie a v rámci nášej koncepcie sme navrhli jednoduchý popis spracovania obrazného jazyka pri poruchách schizofrenického a autistického spektra.

Kľúčové slová: metafora, prirovnanie, kategorizácia, konvencionalizácia, relevancia

Foreword

This diploma thesis concerns with the problem of metaphor comprehension from the viewpoint of psycholinguistics and partially also neurolinguistics and is focused predominantly on the problem whether metaphors are processed as categorizations or comparisons. This work tries to propose a solution to this long lasting problem.

The thesis was written as part of completion of a master study program Cognitive Science offered by the Faculty of Mathematics, Physics and Informatics of the Comenius University in Bratislava.

Contents

Introduction

There are scientific questions which troubled human minds for centuries, and then there are questions people have not cared too much about until last few decades. Scientific research of figurative language and specifically the metaphor is the latter case. Although it would not be correct to say that there was no inquiry into the problem, it certainly wasn't a central issue. As Ortony (1993) notes, for a long time it was usual in scientific and philosophical circles to consider literal language as privileged over figurative language. Overall, figurative language was usually considered to be of aesthetic or rhetoric value, rather than a serious subject for a serious science.

Strikingly, the opposite is true. Metaphors are not only almost omnipresent in natural language, they are also one of the basic tools we use to talk about the world around us (e.g. Lakoff 1992). Metaphors are also used to explain concepts that are not well understood by hearer (Gentner 1982, Stern 2008), often they accompany changes in scientific paradigms (Gentner & Grudin 1985) and it is likely that they facilitate proliferation of new (not only) scientific ideas (Gentner 1982, Gentner & Grudin 1985). In general, metaphors account for a significant part of natural language and thought, therefore they represent an important scientific problem for language sciences.

In metaphor research, one of the main problems is the question whether metaphors are inherently categorizations or comparisons. The categorization view is represented mainly by the so-called *class-inclusion theory* (Glucksberg & Haught 2006). Comparison is partly represented by the *career of metaphor theory* of Bowdle and Gentner (2005). This theory proposes that novel metaphors are processed as comparisons and conventional metaphors are processed as categorizations. Both these theories have some empirical support, but the basic question remains open. We will not concern purely comparison theories (e.g. Ortony 1979), since these were not supported empirically.

In this work, we propose an integrative account using mainly career of metaphor theory (but also valuable insights from class-inclusion theory) and *relevance theory's* account for metaphor (e.g. Wilson & Sperber 2006) which is thought to account mostly for novel rather than conventional metaphors. In this theoretical proposal, we will try to fuzzify the dichotomy of categorization and comparison and try to explain the difference between the two as a continuum. We will also try to explain why sometimes even novel metaphors can be processed as categorizations. Further, we will also try to widen the view and describe,

with only a slight modification, that the same model can account for semantic processing in autism and schizophrenia. We will use Jung-Beeman's (2005) *fine-coarse coding theory* of natural language processing (which is based on *theory of cortical minicolumns*) to account for these disorders.

The thesis is structured in three parts: The first part, represented by chapters 1 to 6, is focused on psycholinguistic theories of metaphor with emphasis on the problem of categorization and comparison. The second part, represented by chapter 7, is devoted mostly to neuro-microanatomical aspects of typical and disordered individuals' brains, with emphasis on the latter. The third part, represented by chapters 8 and 9 consists of a theoretical account integrating the findings from the first two parts with the relevance theory (and some other accounts such as Bayesian approach to word category learning (Tenenbaum & Xu 2000)) and an original study which aimed to test some of the predictions imposed by our integrative account.

Metaphors are intriguing and as the last few decades of scientific research show, study of their dynamic and seemingly odd nature can point us to the very basic cognitive processes such as perception of similarity or hierarchical relationships. We hope that this work will be interesting to a reader with an interest (and at least some preliminary knowledge) in psycholinguistics, neurolinguistics and similar fields.

1 The problem of categorization and comparison

1.1 What is a metaphor?

Usually, a metaphor is defined as an utterance of *X is Y* form, where both arguments are nouns and an utterance is interpreted metaphorically (Stern 2008) – we will use the standard notation and refer to such an utterance as to a *nominal metaphor*. A conventional example of such an utterance is the sentence *My boss is a shark*. The variable *X*, in this case instantiated with *My boss*, is usually called target (alternatively it is also called tenor). Target is what an utterance is about; it is the concept that is being described with an utterance. The other variable, *Y*, here instantiated with the word form *shark* is called base (alternatively it is also called vehicle) and it refers to a concept in terms of which we talk about the target. So the metaphor in general is a trope which allows us to talk about one thing, one concept, in terms of some other concept. We must not forget that when considering form, nominal metaphors, although they are studied most frequently, are not the only type of metaphor – for instance so-called predicative metaphors (e.g. Utsumi & Sakamoto 2010) which metaphorically use a verb, such as *The rumor flew through the office*, are studied as well, although less extensively. In this work, we will focus exclusively on nominal sentences with the impugnable assumption that underlying processes described here should be applicable to predicative sentences as well.

The above mentioned definition fails as a satisfactory definition of metaphor, since it is merely a definition of the linguistic structure studied here, missing the crucial point of metaphor which is its meaning. Saying that a metaphorical utterance is an utterance that is interpreted metaphorically could not be further from being explanatory (since its tautology), therefore we need to look at what is happening on the conceptual level when metaphors are being comprehended or produced. In other words, we need to ask what it means, when considering human cognition, to speak about *X* in terms of *Y*. We should note that from the general communication direction point of view, this work is focused exclusively on metaphor comprehension, not production.

1.2 Categorization and comparison

One of the basic problems in psycholinguistic study of metaphor is the debate about whether metaphorical utterances are inherently understood as comparisons (e.g. Ortony 1979) or categorizations (e.g. Glucksberg & Haught 2006). Both these general views

revolve around the question whether metaphor is semantically equivalent to its respective simile, therefore whether metaphor is elliptical simile or not. Simile is usually defined as a figurative utterance of *X is like Y* form (although there are other forms of simile as well). Simile is obvious comparison, communicating that some X is similar to some Y. In general, comparison theories of metaphor state that metaphor is an elliptical simile, or as Glucksberg & Haught (2006) phrase it in their critique, according to these theories, metaphors are *"understood in terms of their corresponding simile"* (Glucksberg & Haught 2006). This view is traceable far to the ancient Greece, where in the Aristotle's Rhetoric we find the definition of a simile that roughly corresponds to the comparison views of metaphor. Aristotle claims that "*The simile also is a metaphor; the difference is but slight*"(Aristotle 2004). To demonstrate this, we can look at the examples from Rhetoric, rephrased by Stern (2008):

(1) Achilles is a lion.
(2) Achilles is like a lion.

It is intuitively clear that between (1) and (2) the difference, if any, is truly *but slight*. It doesn't seem to be true to say that there is a large semantic shift between the two forms. With respect to this, there is a notion claiming that the difference between metaphor and simile is in the relative strength of content being communicated, with metaphor being stronger than simile (Chiappe & Kennedy 2000). Consider this example fragment of a dialogue:

(3) A: "Achilles is like a lion."
B: "No, he is a lion."

In (3), the speaker B corrects the view of the speaker A, who is uttering a simile, with stronger, metaphorical expression. This is referred to as *correction convention*, and indeed, it seems to be nothing more than just a convention, since when a metaphor or its respective simile appear independently in the same context, this effect of metaphor being stronger than simile basically disappears (Chiappe & Kennedy 2000).

Based on these few simple arguments, it might seem that the basic problem of metaphor is solved and we can describe metaphor simply as elliptical simile. Problems arise when we look at the difference between the literal statements of the same forms as metaphor and simile, such as:

(4) Dog is an animal.

(5) Dog is like an animal.

In cases (4) and (5) it is quite clear that ellipsis doesn't hold. *Dog is an animal* means that dog is a member of the category of animals, but saying that dog is *like* an animal is semantically incorrect, since this imposes resemblance relationship instead of categorical assertion.

Examples like these prepared ground for the categorization view of metaphor. In the categorization view, represented here by the so-called *class-inclusion theory* (Glucksberg 2003, Glucksberg & Haught 2006), there is a significant difference between metaphor and simile. Class-inclusion theory asserts that when an utterance such as *My lawyer is a shark* is uttered it is always a categorical statement. This means that the base term does not refer to the literal category *SHARK*[1], but to the metaphorical category *SHARK**. This metaphorical category *SHARK** is different from the literal category *SHARK* from which it was originally derived in such way that the first (the metaphorical) is hypernym of the latter. *SHARK** is the category that contains both, lawyers and literal sharks, the category with properties such as being vicious, merciless or aggressive, but not necessarily being able to swim or having sharp teeth (Glucksberg 2003). This function of metaphorical base terms allowing them to refer to both literal and abstract metaphorical category is called *dual reference*.

Class-inclusion theory treats similes as assertions of similitude and therefore they are thought to be processed differently and to have different meaning than metaphors (Glucksberg & Haught 2006). However, the view of metaphors as categorizations is easily acceptable if we consider conventional metaphors only. When considering novel metaphors, the base of a metaphorical sentence doesn't refer to an appropriate metaphorical category prior to utterance's production. Such category must therefore be created *ad hoc.* It is, however, questionable whether we can describe novel metaphor processing as categorization instead of probably simpler comparison (see chapter 3).

On behalf of the categorization view speaks also the fact that the mentioned impossibility to rephrase literal sentences from categorization form (X is Y) to comparison form (X is like Y) (as in (4) and (5)) is not exclusive to literal language. There are metaphors which,

[1] When necessary, we will use italicized capitals to refer to concepts. Asterisk is used to distinguish different concepts with the same word form.

although they are novel, seem to be semantically odd (e.g. (6)), or are simply incorrect (e.g. (7)) in comparison form, such as these Glucksberg's & Haught's (2006) examples:

(6) Florida is headed to be (like) the next Florida.
(7) My boss is (like) a well-paid shark.

Besides not being very aptly rephrasable to comparison form, the sentence (6) uses the word *Florida* in both senses, the superordinate metaphorical (base position) as well as in the subordinate literal (target position); therefore the categorization view of this sentence's processing seems to be necessary (Glucksberg & Haught 2006). Sentence (7) does not seem to be meaningful in comparison form at all and we will analyze such type of sentences in chapter 3.

Of course, categorization as being opposed to comparison is a rather odd notion, since to categorize something, it is essential to compare the properties of the concept being categorized with the categorical concept (Barnden 2012), unless the category is defined by listing its members, but that is, of course, not the case here. Anyway, the class-inclusion theory of metaphor has many advantages, maybe the most notable being its straightforward account for metaphor directionality. The directionality of metaphor is one of the basic motivations for the above described terms of target and base. Consider these examples from the study of Bowdle & Gentner (1997):

(8) Life is a journey.
(9) Journey is a life.

It is quite clear that (8) is just a conventional metaphor, but the sentence (9) is, if not nonsensical, then very hard to interpret. The point is that if we take two concepts, there is usually a preference for one or the other direction, in other words, for specific instantiation of base and target. Class-inclusion theory easily accounts for the directionality of metaphors, since the base position is always instantiated by the superordinate category, member of which is the target. However, this is not the only plausible account for the metaphor directionality and we will return also to this issue again in chapter 3.

Both comparison and categorization theories have some theoretical ground that cannot be ignored. Based on this, our further investigation into the problem of metaphors will be based mainly on the so-called *career of metaphor theory* (Bowdle & Gentner 2005). In the light of the previously mentioned views of metaphor, this theory can be seen as integrative.

In heart, the career of metaphor theory can be considered a comparison theory, but it states that during the course of metaphor learning there is a switch of processing type from comparison to categorization. For instance if the utterance *My lawyer is a shark* is encountered for the first time, the comprehension process is based on comparison. As the metaphor is frequently used with the same or very similar interpretations of the base term, the abstract concept that is created as byproduct of the comparison process becomes memorized as conventional interpretation of the base term. This process is called conventionalization of metaphor and in the career of metaphor theory it is considered to be the main variable responsible for the switch of the processing type. Comparison and the process of conventionalization in the career of metaphor theory are well defined and are based primarily on *structure mapping theory* of (not only) analogy, which will be presented in the next chapter.

2 Structure mapping theory

2.1 Relations, attributes and metaphors

Structure mapping theory (Gentner 1983) is a theory that represents the core of career of metaphor theory. It provides a simple yet effective ontology that helps us to categorize different types of similarity based on the number and type of their properties and also to describes the process of emergence of similarity (the comparison process) when considering conceptual structures. Concepts or knowledge in general is represented here using a *„propositional network of nodes and predicates"* (Gentner 1983), where *"the nodes represent concepts treated as wholes; the predicates applied to the nodes express propositions about the concepts"* (Gentner 1983). In the structure mapping theory, two basic proposition types that show significant qualitative differences are distinguished. First are attributes which are predicates with one argument only, such as property of being red RED(x)[2]. The other type are relations, standardly defined as predicates with two or more arguments for instance collision COLLIDE(x, y). When we talk about analogy in structure mapping theory, we talk about similarity between two relational structures. For instance, in some sentence with structure *A is analogical to B* or simply *A is (like) B* that is an analogy, attributes are not important and the only the part of the two propositional networks that matters are their relations. Consider this classic example of analogy used repeatedly by Gentner (e.g. Gentner 1982, 1983):

(10) The atom is like our solar system.

To explain the analogy in (10), Gentner uses the classical Rutherford's model of the atom and the set of correspondences between the concepts, encoded by the sentence, is probably immediately obvious to the reader. There is a core in the atom around which the electrons revolve which is analogical to the sun around which the planets revolve, the electrons are smaller than the nucleus and similarly planets are smaller than the sun, the core attracts the electrons and not vice versa and the situation is similar with the sun and planets. Features of concepts of the atom and the solar system can all be represented as relations such as REVOLVES AROUND(*planet*, *sun*), REVOLVES AROUND(*electron*, *nucleus*). Most attributes don't match and those that match are not important. For instance, the core of an atom is not millions of degrees hot, but this fact has no effect on the appropriateness of this

[2] Non-italicized capitals refer to predicates, in accord with Gentner (1982, 1983).

analogy. Some attributes match, for example, nucleus, electrons, the sun and planets, all are usually imagined to be more or less round – but if, for instance, the core the atom would be cube-shaped it would not make the analogy less appropriate because the relational similarity would remain unchanged. Gentner & Markman (1997) termed this characteristic property of analogy as *relational focus*.

To summarize, two concepts or conceptual domains are analogical when they share relational structure, and attributional similarity is nonexistent or is not important (Gentner 1983, Gentner & Markman 1997). On the other hand, when talking about metaphors, the situation is more diverse – many or most metaphors seem to rely heavily on relations (Zharikov & Gentner 2002), but there are also metaphors that rely exclusively on attributes, such as this example sentence from Zharikov & Gentner (2002):

(11) The cloud is like a marshmallow.

Authors hypothesize that these *attributional metaphors* usually don't encode something interesting or needed and therefore they are less likely to become conventionalized (Zharikov & Gentner 2002). As Zharikov & Gentner (2002) note about the sentence (11),*"we already have words for white and fluffy"* (Zharikov & Gentner 2002), hence the use of such figurative statements doesn't enrich the communicated content and therefore the speaker is less likely to utter them. It is also easily conceivable that it is harder for a hearer to decode such sentence than just the list of two well-known attributes (*The cloud is white and fluffy*) and therefore such utterance is suboptimal for this reason as well. On the other side, there also may be contexts, although possibly less frequent, where it is the connotation that is pronounced and primarily communicated, and even such an attributional utterance as (11) becomes optimally relevant.

Presented distinction between attributes and relations proves to be very effective since it allows a description of various types of comparison within a single framework. We can identify: identity - where both, all attributes and all relations are shared; literal similarity - where a large number of both; attributes and relations is shared, analogy - where only relations are shared (such as (10)); appearance match - where only attributes are shared (such as (11)); and sentences that are incomprehensible or anomalous that share neither or very small amount of properties. Metaphors, in contrast with these various comparison types, represent a rather unspecific case since they are not well definable using parameters

used by the structure mapping theory, and as was said earlier, they can be relational or attributional, but also a combination of both (Gentner & Markman 1997).

What is interesting about this ontology is that it doesn't provide us with some discrete separation of literal and figurative comparison types but rather proposes a continuum based on the number and type of shared features.

What was mentioned in this chapter so far allows us to differentiate between the two basic types of properties that mental objects can have and shows us that there are different ways in which any two thinkable objects are or are not alike, given the extent and type of structural similarity. This alone provides us with the universe of potential comparisons, but does not explain how these various types of comparison come to existence. Mechanisms through which these mappings of structures are achieved are described in the next section.

2.2 Structure mapping mechanisms

Two basic mechanisms of the structure mapping process are recognized in this theory and these are *structural alignment* and *projection* (e.g. Gentner & Markman 1997, Gentner et al. 2001, Bowdle & Gentner 2005).

2.2.1 Structural alignment

Propositional structures are first aligned. Alignment, in order to be successful must satisfy the two main constraints – structures that are mapped onto each other must be *structurally consistent* and *systematic*:

a) Structural consistency depends on two conditions. First, aligned concepts must show *parallel connectivity* which means that when two aligned concepts have matching predicates they also must have matching arguments of these predicates. For instance, in the above mentioned example predicate REVOLVE AROUND(x, y) matching between the concepts *SOLAR SYSTEM* and *ATOM* both arguments are matching as well – the sun is equivalent to atom's core and planet is equivalent to electron.
   Second condition is the so-called *one-to-one mapping – „where any element in one representation is limited to at most one matching element in the other representation"* (Gentner & Marksman 1997). This means that for two concepts, to be structurally consistent, parts of their structures must be unambiguously similar. For instance, we cannot match atom's core with both planet and the sun.

b) Systematicity reflects a preference for matching of higher order relations (relations of relations; e.g. second order relation is a relation which arguments are first order relations) – meaning that if there are multiple potential alignments of two concepts with an equal number of matching elements, the one with higher order relations will (everything else equal) always be preferred. This preference for deep relational structures shows that although it is true that *"any two things can be alike in any number of ways"* (Glucksberg & Haught 2006), rough quantity of overlapping features (Tversky 1977) or their salience (Ortony 1979) are not the only valid criterions. For example, if the first would be true, it would be fine to compare *slavery* with *orange*, since they both are possible to exist, they both are thinkable and so on. As Gentner & Markman (1997) informally note, such comparisons make people confused or angry.

With the account for systematicity we can return to the problem of metaphor directionality, the preference for concepts to specifically instantiate the base or the target position, which we briefly mentioned in previous chapter. Bowdle & Gentner (1997) in a series of experiments used pairs of short stories and found that it is systematicity that drives the preference for certain direction of comparison of these stories, with more systematic story being preferred on the base position. Important is that with this notion, we can leave the view that directionality is driven *necessarily* by the hypernym being on the base position, and hence also leave less space for categorization-based theories of metaphor (Bowdle & Gentner 1997). However, it may be still possible that a more systematic concept is perceived as a hypernym, due to systematicity.

There are also other, but weaker accounts for directionality of metaphors, for instance, conceptual metaphor theorist[3] Kövecses (2010) reports that bases are usually more concrete and targets more abstract. This is a valid and truly important observation, but Kövecses doesn't explain why this is so, although Lakoff (1993) notes that metaphor targets are less structured than metaphor bases and therefore the basic idea, although less elaborated and proposed exclusively for metaphors, is similar to the one of systematicity. Structure mapping theory therefore can easily account for mostly primary metaphors[4] (e.g. Gibbs, Lima & Francozo 2004) described by the above mentioned authors with the simple assumption that concrete sensory-motor concepts are indeed, for obvious reasons, more systematic.

[3] We will talk about the conceptual metaphor theory in chapter 4.
[4] Metaphors mapping sensory-motor concepts with subjective experience (e.g. *GOOD is UP*).

In her early work, Gentner (1982) also identified analogies which violate conditions of the structural consistency and express low systematicity, yet they are not considered meaningless. She termed these *expressive analogies* as opposed to *explanatory analogies* (such as scientific analogies and basically all analogies and metaphors we did and will mention). She identified these expressive analogies/metaphors in poetry and besides their deficits in the above mentioned structural consistency, they also showed increased *richness* (quantity of nodes mapped) (Gentner 1982). It is questionable whether we should consider these expressive, sometimes also termed *poetic metaphors* a separate kind or type of metaphor. I won't further elaborate on this issue in other way than by stating my doubts about poetic metaphor using Stern's words: *"I'm not sure, whether such a beast exists"* (Stern 2008).

2.2.2 Projection

When two structures are aligned, projection from the base to target domain occurs – not only initially aligned elements are projected, but more importantly, new elements from the base domain are also projected as so-called *candidate inferences*. Some candidate inferences, of course, prove to be valid (consistent with the target concept) and some does not (Gentner 1982). This proposition of the acquisition of an element from the base is very effective since it can account for two (however, inherently indistinguishable) important aspects of metaphor. First aspect is the remarkable ability of some metaphors to clarify a target concept that is not well understood or defined. Some parts of concepts are aligned at first and then through candidate inferences that prove to be valid, hearer can elaborate the structure of a target concept. Notice that also the previously mentioned notion of systematicity being the driving force of the metaphor directionality makes a perfect ground for this explanatory use of metaphors, since it is naturally presumed that more systematic concept is used to describe some other not well structured concept and not the opposite (Bowdle & Gentner 1997). Secondly, candidate inferences can account for so-called *extended mappings* between conceptual domains (Gentner et al. 2001) or *conceptual metaphors* as Lakoff (1993) terms them. We often observe metaphors to be used not only as isolated comparisons – quite the opposite; some are seen to spread consistently over many situations and sentences. Take for instance Lakoff's notoriously known example of such extended mapping (or conceptual mapping or, in his conception, conceptual metaphor) - *LOVE is A JOURNEY*. This mapping is reflected in many metaphorical expressions such as: *Look how far we've gone, We can't turn back now, The marriage is*

on the rocks, The relationship isn't going anywhere and many others (Lakoff 1993). Question of extended mappings is important for the understanding of the metaphor and analogy, and we will return to this topic in chapter 4.

In the next chapter we will finally talk about the career of metaphor theory and approach some of the problems mentioned in chapter 1.

3 Career of metaphor theory

The career of metaphor theory is an account that directly builds on the structure mapping theory. As already mentioned, this theory introduces the notion of metaphor conventionalization in terms of different processing of novel and conventional metaphors. Initially, when a metaphor is encountered for the first time[5], the metaphor is processed as a comparison of two literal concepts. Bowdle & Gentner (2005) claim that when a metaphor is novel, new abstract metaphorical categories are mere byproducts of the figurative comparison process, and therefore they don't have any influence on interpretation of these comparisons. With repeated use that leads to similar interpretations of a base term, a new metaphorical category is derived which now a base term conventionally refers to. Now that this metaphorical category is entrenched in conceptual system it is more likely to be approached than an original literal concept when interpreting metaphors, since it is less computationally costly to use an existing category than to derive an interpretation from scratch every time a metaphor is encountered (Bowdle & Gentner 2005). Conventional bases therefore always have the previously mentioned *dual reference* function – they can be used to refer to the original literal concept or to the abstract superordinate concept/category, member of which is also the original literal concept.

Because according to this theory, metaphors are initially processed as comparisons, people should prefer comparison form over categorization form whenever they process a novel metaphor. The opposite, preference of categorization form over comparison form should hold for conventional metaphors. This hypothesis was repeatedly tested in series of experiments (Bowdle & Gentner 1999, 2005) – when participants were instructed to choose between the comparison and the categorization form versions of the same metaphor, they were more likely to choose the comparison form when metaphor was novel and the categorization form when metaphor was conventional. Further, they observed that novel metaphors are processed faster in comparison form and conventional metaphors in categorization form. Third and truly beautiful experiment of Bowdle & Gentner (1999, 2005) used a method they termed *in vitro conventionalization*. Participants tested with this method were presented with triplets of sentences in the comparison form and with similar base terms. The first two of these sentences were complete, meaning that they had their targets instantiated, but the third sentence had blank space instead of a target. Participants

[5] Saying that some metaphor is novel doesn't necessarily mean that the metaphor is encountered for the first time; however, we will use the notion of a first encounter as a model situation of novelty.

had to create a new appropriate target to fill in the blank position with, based on the meaning of the first two sentences. An example of this procedure (Bowdle & Gentner 1999):

a. An acrobat is like a butterfly.
b. A figure skater is like a butterfly.
c. _____ is like a butterfly.

After this conventionalization phase, participants were presented with various metaphors in comparison and categorization form, some of which had similar base and meaning as the sentences in the conventionalization phase and were asked to rate their preference for one form over the other. The results were in accord with the career of metaphor theory – sentences used in the conventionalization phase, were preferred in categorization form, but novel metaphors that did not, were preferred in comparison form (Bowdle & Gentner 1999).

Another study by Wolff & Gentner (1992) which tried to prove the proposal of career of metaphor theory used the priming paradigm. In this study, participants had to paraphrase metaphorical sentences that were primed by either their base or target. Authors successfully showed that when a metaphor is novel, there is no difference in the onset of response regardless of the prime condition, but when a metaphor is conventional the onset of response is significantly lower when the prime word is base than when the prime word is target. Interpretation of these results supports the career of metaphor theory – in conventional metaphors, there already is a metaphorical category present in the conceptual system, and therefore categorization occurs. In novel metaphors, base and target concepts are initially equivalent, since the match between their structures must be found first, therefore these results support the comparison view of novel metaphors (Wolff & Gentner 1992). This notion may seem contradictory with previously mentioned account for directionality, but it is not since the career of metaphor theory assumes that comparisons are role-neutral during structural alignment and directionality arises in the projection phase (e.g. Gentner et al. 2001).

Valid question concerning these experiments and career of metaphor theory is, whether these effects of novelty and conventionality are universal ones, and whether they apply to all metaphors. As Glucksberg and Haught (2006) show, we can identify and easily construct metaphors that are novel, but more apt in categorization than in comparison

form. Example of such a metaphor was already presented in (7). Metaphors that are not apt in comparison form can be constructed if we modify the base with an adjective that is applicable to target concept only. (12) and (13) show original unmodified conventional metaphor and metaphor modified with target applicable adjective, respectively (Glucksberg 2006):

(12) Some ideas are (like) diamonds.
(13) Some ideas are (like) *theoretical* diamonds.

Obviously, being theoretical is not a common property of literal diamonds. Pseudo-novel sentences such as (13) are preferred in categorization over comparison form, since they seem to be incorrect in comparison form and seem to directly imply categorical assertion. Although sentences such as (13) may represent relatively infrequent uses of figurative language, their existence itself seems to be an important phenomenon which a comprehensive theory of metaphor must explain. Glucksberg & Haught (2006) tested this observation experimentally. Besides control unmodified sentences, three types of stimuli sentence modifications were used – target applicable adjectival modifications such as (13), base applicable adjectival modifications (e.g. *some ideas are (like) gem-cut diamonds*), both target and base applicable adjectival modifications (e.g. *some ideas are (like) small diamonds*). In comparison form, sentences with only target applicable modification of base, took longer to process (longer comprehension time) and scored lower in aptness and comprehensibility ratings than all other modified sentence types. These results were replicated on genuinely novel metaphors as well (Glucksberg & Haught 2006). This outlying difference seems to be a sign of rather anomalous sentences, such as the one in (5) – dog cannot be like an animal in the same sense in which idea cannot be like a theoretical diamond. The major problem with this argument against the comparison view of novel metaphors is that *theoretical diamonds* is adjectival metaphor itself. Therefore in sentence (13) the target concept is compared to a new metaphor. We may term this a *second order metaphor*, since it encompasses metaphorical processing at the level of the modified base as well as at level of the whole nominal sentence. We should therefore ask, whether this says more metaphors or about flawed methodological confusion of metaphor for any nominal metaphor.

An interesting view on the problem of selection of novel metaphors in the categorization form but also on conventionalization in general can be drawn from the study of Zharikov

& Gentner (2002). In their research, authors presented participants with short paragraphs which served as targets. For every paragraph, participants made forced decision whether they preferred a metaphorical sentence in comparison form or in categorization form as a description of a paragraph. The bases varied in conventionality, and more importantly, paragraphs varied in their dominant predicate type – they were either attributional or relational. Interestingly, the study showed that even though the comparison form was more often preferred for both novel attributional and novel relational paragraphs, when categorization form was preferred for novel metaphor, it was significantly more often when the paragraph being described was relational.

These results are in accord with Aisenman's *relational precedence hypothesis* (1999) which suggests that categorization form invites relational interpretation and comparison form invites attributional interpretation and also with career of metaphor theory. Based on this, we can say that it is possible that when a speaker wants to emphasize relational aspects of an utterance, he is more likely to use categorization form, even though a metaphor is novel. A plausible explanation of this phenomenon may be found in the study of Gentner & Marksman (1997), where authors showed that people are in general more prone to process relations than attributes and therefore it may be easier to derive a new abstract category when interpretation of an utterance is relational. This would suggest that sometimes, in disagreement with the career of metaphor theory, new categories can be derived and readily used during the comprehension of novel metaphors.

4 Extended mappings

Although the *X is (like) Y* sentence form is more than often used in studies of metaphors, isolated utterances of this form are far from all that there is to metaphors and their extensive use in common language. The theory which is concerned predominantly with these extended or large-scale mappings is the previously mentioned conceptual metaphor theory (e.g. Lakoff 1993). An example of one such mapping, *LOVE is A JOUNREY* was mentioned at the end of chapter 2. There are many such extended mappings listed by conceptual metaphor theorists, such as *ARGUMENT is WAR, THEORIES are BUILDINGS, IDEAS are FOOD* and many others.

According to the conceptual metaphor theory, extended mappings represent a substantial organizational principle of our conceptual system. Conceptual mappings are considered here to be so significant that this theory proposes *online comprehension* only. Online comprehension of metaphor is a theoretical construct that suggests that every time a metaphor is processed, both conceptual domains (base and target) are activated, since the base domain in this theory provides organizational and linguistic structure to the target domain. That means that when someone hears about *LOVE*, he needs to active the domain *JOURNEY* to make understanding possible. Opposite to online comprehension is so-called *offline comprehension* which proposes that metaphorical targets may be understood without the need for base domain activation. There are at least two studies testing this proposal that large-scale metaphors are comprehended online. Both Gentner & Boronat (1992 as cited in Gentner et al. 2001) and Keysar et al. (2000) reported similar conclusions, claiming that extended mappings (in terms of online comprehension) are used only during comprehension novel metaphors, but not conventional metaphors.

Gentner & Boronat observed how reading of novel or conventional metaphors is influenced by preceding context. In this study, metaphors were preceded with metaphorical contexts that were either consistent or inconsistent with an observed metaphorical sentence. Preceding literal context served as a control condition. Novel mapping was for instance *A DEBATE is A RACE*; conventional mapping was for instance *A DEBATE is A JOURNEY*. The authors reported that when a metaphor was novel, it benefited from preceding context and faster reading times were observed, but when a metaphor was conventional there was no effect of preceding context and therefore there was no difference in reading times. This seems to be sufficient evidence to reject the claim of the

conceptual metaphor theory about online comprehension to be universal. Gentner et al. (2001) hypothesized that conventional metaphors are processed in a localist manner, with lexically stored interpretations, which is in accord with notion of the career of metaphor theory that conventional metaphors are processed as categorizations. Because conventional metaphorical bases are well learned categories, there is no need to approach the original literal base concept and search in a large-scale network. Indeed, the fact that humans don't use large-scale mappings when they utter conventional metaphors, yet they can draw new inferences about metaphorical concepts based on their literal counterparts (or invent theories of metaphor) is a nice proof that they can re-approach original literal concept, even when there is a well established metaphorical category.

Study of Keysar et al. (2000) also examined reading times of metaphors preceded with meaningful conventional and unconventional contexts. In the first experiment, they found, similarly to the previous experiment, that when metaphors are conventional, there is no benefit for metaphor reading times whether they are preceded with explicitly or implicitly presented conventional metaphorical context or literal context. On the other side, in the second experiment, when the same metaphors were preceded by unconventional context (words conveying metaphorical meaning were changed for unconventional synonyms), they benefited from the extended mappings and were comprehended significantly faster than metaphors in other conditions. Interesting about this study is that metaphors were not even truly novel, only the context was rephrased – this shows that interpretations of conventional metaphorical bases are strongly connected to their lexical forms.

In conclusion, it seems that the conceptual metaphor theory does not account for extended mappings very well, and the proposal of the career of metaphor theory and its proposal of candidate inferences as part of structure mapping process fits the data more comprehensively.

5 Dead metaphors

So far, we have talked exclusively about novel and conventional metaphors, but there is also a third stadium we can identify through the course of metaphor „development“ and that is its death. *Dead metaphor* is a privative, denoting a term that used to be metaphorical but it is not anymore (Stern 2008). A common example of an English dead metaphor is the word *blockbuster* that is used to refer to a very popular and successful movie. Originally this word used to refer to a strong explosive that could demolish an entire city block but the word has lost this use in everyday language. Another example of a dead metaphor is the word *culture* that can refer to both a human culture as well as a bacterial culture. Importantly, meanings of these words or the concepts that the words refers to are not directly related anymore and therefore the concepts are referred to independently – they are basically homonymous.

Bowdle & Gentner (2005) on the base of these examples distinguish two kinds of dead metaphors – dead metaphors the original literal term of which they were derived from is still actively used (such as *culture*) and dead metaphors where only the derived meaning remains in common language use (such as *blockbuster*). Interesting for our investigation is that this indirectly shows that during the course of metaphor conventionalization, an abstract base concept is derived and stored in memory. The fact that this abstract concept is stored in the memory seems obvious but still, this notion is sometimes omitted. For instance the relevance theory (Wilson & Sperber 2006) postulates that metaphorical use of a literal word is always an *ad hoc* concept created by process of lexical loosening (we will briefly describe the notion of lexical loosening later in this chapter and more thoroughly in chapter 8).

Another important aspect of metaphor which the notion of dead metaphors points to is that metaphors don’t just allow us to speak about one thing in terms of some other thing; they also have a significant role in the process of introduction of new literal concepts to natural language[6]. This, as showed above, happens when a metaphor dies, and it is probably reasonable to think that a metaphor is usually conventional before it dies, otherwise the abstract meaning would die with the literal meaning.

[6] For more radical view on this matter, consider Nietszches notion that all literal language has its origins in metaphor (e.g. Hossain 2013).

Another opinion on this matter we want to mention concerns the nature of difference between live and dead metaphors as introduced by Stern (2008). Although he recognizes multiple types of metaphorical vitality or degrees of liveness, he emphasizes that *„the more the interpretation of a metaphor Ø in a context c depends specifically on c, i.e., on presupposition associated with Ø specific to that context c, the livelier the interpretation of the metaphor"* (Stern 2008). Stern's view posits that it is the context dependency of metaphorical interpretation of a base concept that is central to the concept of metaphor vitality.

Although the whole Stern's conception is semantic and not pragmatic, it seems to account for context dependency in a wider sense than linguistic semantics usually does. The main problem with Stern's view of metaphor vitality is that it doesn't distinguish lexical forms and concepts. Consider, for instance, these metaphorical uses of the word *sun* that Stern frequently uses:

(14) Juliet is the sun.
(15) Achilles is the sun.

It is clear, that (14) and (15) are used here to communicate rather different properties. The first sentence is from the Shakespeare's Romeo communicating that Juliet is the center of his world; his day begins with her rising and so on. On the other hand we see the description of Achilles as angry and powerful force of the nature. The *sun* in these sentences communicates disjoint sets of properties, both derived from the original literal concept of the sun in a context dependent manner.

There are at least two reasons to consider this view questionable. First, if we distinguish concepts from word forms, we can easily account for situations such as (14) and (15) in accord with career of metaphor theory. If the concepts referred to by the word *sun* are conventional, the metaphor vitality in Stern's conception is equal to polysemy or homonymy and hence the above cases are not more context depended than any literal polysemy or homonymy. By implication, this notion would classify all conventional metaphors as dead which would be an unnecessary terminological shift.

If a metaphor is novel, only then we can talk about context dependency in metaphor-specific way, because the metaphorical interpretation is being directly derived. Also if a conventional base term is used in a novel way with novel interpretation, it should be

considered a novel metaphor and not a context dependent use of a conventional metaphor since it requires a new comparison or an *ad hoc* category creation, based probably on the original literal, rather than abstract metaphorical concept, especially when the properties communicated by a novel metaphorical interpretation are disjoint from the properties communicated by the already existing conventional metaphorical interpretation.

The second reason is represented by now classical pragmatic accounts for numerous cases of context dependent interpretations of literal terms (e.g. Dolník 2013, Glucksberg 2001), the notion that there is no strong base for the dichotomy of literal and figurative language (e.g. Glucksberg 2001) and the relevance theoretic (Wilson & Sperber 2006) account of *lexical loosening* and *narrowing* which shows that even the most ordinary literal words such as *bank*, can have numerous different interpretations in different contexts. The word *bank* can refer to some bank institution in general, to some specific bank (narrowing), or to any institution that provides money under similar conditions, but is not a bank by the definition (loosening). These almost unnoticeable context-dependent interpretations are very common in natural language, and according to Wilson & Sperber (2006) they are governed by the same principles that govern interpretation of metaphors. Metaphor here is just a more extreme case of lexical loosening and therefore context dependency alone does not seem to account for metaphor vitality sufficiently.

The above mentioned proposition of Bowdle & Gentner (2005) seems to be more plausible than Stern's one. It also implies, that perception of metaphoricity or metaphorical vitality is caused merely by our ability to identify and possibly rederive a metaphorical concept from its original literal concept, and therefore blurs the difference between literal and figurative language even more than it already is.

6 Aptness, conventionality and possible methodical issues

As mentioned previously, the career of metaphor theory claims that the main variable responsible for the difference between comparison and categorization processing type is the process of conventionalization. On the other side, Glucksberg's class-inclusion theory claims that the main variable is the aptness. Both these claims seem to have some empirical ground. As mentioned previously, Bowdle & Gentner (2005) found that comparison form of metaphor is usually preferred when the metaphor is novel and categorization form is preferred when the metaphor in conventional and also that novel metaphors are processed faster in comparison form and conventional metaphors in categorization form. However, Glucksberg & Haught (2006) analyzed Bowdle's & Gentner's data and found significant correlation of aptness and conventionality but also of aptness and conventionality. On this ground, they described the results of Bowdle & Gentner as confounded with aptness since it is not possible to determine whether and to what extent their results are due to effect of conventionality or aptness. Other studies (e.g. Pierce & Chiappe 2009, Jones & Estes 2006) that tried to disentangle the problem of aptness and conventionality, usually reported results opposite to the Bowdle's & Gentner's – the difference between comparison and categorization form preference was directed by metaphors' aptness in such way that metaphors with low aptness ratings were preferred in comparison form and metaphors with high aptness ratings were preferred in categorization form and there was no effect of conventionality.

However, a recent study by Thibodeau & Durgin (2011) showed that the method usually (and in all above mentioned studies including Bowdle & Gentner (2005)) used to assess conventionality may be flawed. This method determines conventionality of a metaphor as a function of the base conventionality which is determined by subjective ratings of how common it is for the base term to communicate a certain property. A typical example of this method's rating task is: *„How common is it to use the word 'blueprints' to refer to something that 'provides detailed instructions on how to build something'?"* (Pierce & Chiappe 2009). Thibodeau & Durgin (2011) found that conventionality ratings assessed using this method, don't correlate with conventionality ratings of whole metaphorical sentences. This makes sense, since if a base conventionality would be the only determiner of a sentence conventionality, all figurative sentences with the same base but different target would be equally conventional. This would imply that for instance the sentence *My*

lawyer is a shark would be as conventional as the sentence *Slavery is a shark* (Thibodeau & Durgin 2011).

When Thibodeau & Durgin (2011) measured and analyzed the conventionality ratings of whole sentences, they found that they strongly correlated with the aptness ratings of these sentences. They also found that this metaphor-sentence conventionality correlates with the frequency of a sentences' occurrence in the corpus which they consider to be the most objective method to assess conventionality. Metaphor-base conventionality ratings were not shown to correlate with corpus frequency and they were also shown to be independent of aptness ratings, which is in accord with the critique of studies using metaphor-base conventionality assessment and considering metaphor conventionality a variable independent of aptness.

Further, Thibodeau & Durgin (2011) also found problems with the standard assessment of metaphor aptness which they found to be enhanced with in vitro conventionalization, with the frequency of sentences in corpus, but also with the defining property of aptness, the number of important base features communicated or not communicated about the target. Authors argue that the aptness ratings capture processing fluency or ease of processing and sometimes even conventionality, rather than the quality of the metaphor itself (Thibodeau & Durgin 2011).

Based on this critique, it is rather hard to assess what exactly which study measured. If authors used metaphor-base conventionality measurements, but also used them with their conventional targets, results of these studies might be fine, but as Thibodeau & Durgin (2011) demonstrated, this certainly is not always so, since in their study, stimuli from the study of Jones & Estes (2006) were used, and therefore at least in this case the critique was appropriate.

However, as in vitro conventionalization, that was used in this study as well, shows, there is an apparent effect of conventionalization that seems not to be confounded by the above mentioned problems. Authors found that in vitro conventionalization enhances processing time, which is consistent with the original results of Bowdle & Gentner (1999, 2005) and as already mentioned, enhances aptness ratings. Based on this, authors conclude that *"conventionalization certainly plays a role in determining the speed and ease of metaphor processing"* (Thibodeau & Durgin 2011).

Regarding the concept of aptness, we think that it is highly doubtful whether we can talk about conventional metaphors that would not be apt. It would be quite odd if a metaphor would become conventional but not apt to communicate some features. It is questionable, then, whether scientists are asking the right question when they are trying to define aptness and conventionality as independent dimensions. In general, we think that it is not possible for conventional metaphors not to be apt and that aptness can vary significantly in novel metaphors only.

7 Brain, minicolumns and metaphors

Although there is a lot of neuroscientific research of natural language comprehension, this research is, for obvious reasons, usually focused on high level neuroimaging techniques such as EEG or MRI. It might seem odd that in metaphor research which tries to describe a rather high level phenomenon (although in previous parts it was showed that metaphor comprehension encompasses the very basic operations with concepts) one of the strong theories is based on findings from the field of microanatomy. The so-called *fine-coarse coding theory* proposed by Jung-Beeman (2005) is based on the *theory of cortical minicolumns* (e.g. Hutsler & Galuske 2003). In this chapter, we will first introduce the minicolumn theory in general and then in the context of metaphors. Importantly, we will also try to explain metaphor comprehension in autism and schizophrenia spectrum disorders using the same paradigm. An account for semantic processing in autism and schizophrenia can be considered the main motivation of this chapter, since its results will be used in chapter 8, where we will propose a theoretical integration of theories mentioned in previous chapters and the relevance theory and try to show that it can easily be modified to account for semantic processing observed in these disorders. However, for a reader who is not interested in neural correlates of figurative language processing, it is fine to skip this chapter.

7.1 What is a cortical minicolumn?

Minicolumns, by some considered to be the smallest processing units of the brain (Hutsler & Galuske 2003), can be defined as vertical alignments of pyramidal neurons. Vertically these minicolumns are 3-4 mm long (Chance 2014), therefore they span through all cortical layers. Horizontal width (further just 'width') is more crucial for our investigation and Shenker et al. (2007) formulate its definition as *„a measure of the size of the core region of the minicolumn, which contains the majority of the neurons and apical dendrites, and both myelinated and unmyelinated fibers"* (Schenker et al. 2007). Width is standardly observed to vary with species, hemisphere, and cortical area, but also with psychopathological conditions such as schizophrenia, autism, dyslexia and others. Although there is a lot of variation, an average minicolumnar width is approximately 60-90 μm (Buxhoeveden & Casanova 2002 as cited in Hutsler & Galuske 2003). Additionally, we can also identify higher-order organizational structures – macrocolumns, which consist of minicolumns and are approximately 600-800 μm wide (Hutsler & Galuske 2003).

Functionally, it was found that cells inside a minicolumn have more similar receptive fields with each other, than with cells from adjacent minicolumns (Opris et al. 2012 as cited in Chance 2014) – this observation provides a support for the notion of minicolumns as basic processing units, or we could better say processing-organizing units, of the brain.

7.2 Minicolumns and hemisphericity

Numerous cognitive faculties show hemispheric specialization, for instance language comprehension and production is (in general) dominantly processed in the left hemisphere whereas for instance face recognition shows a right hemispheric advantage. Since these specializations are constant in neurotypical population, there obviously are some constant anatomical and physiological causes of these hemispheric asymmetries. Sometimes the size of a region serves as an anatomical indicator of hemispheric specialization, but as Chance (2014) mentions there are exceptions. With this in mind, it seems to be promising to look at the microanatomical level and try to infer what kind of properties a neuronal network must have to process some type of information more efficiently than other. The main variable we will be interested in is the spacing between minicolumns which will be referred to as minicolumnar spacing.

Typical example of both functional and structural hemispheric asymmetry is the planum temporale (PT), an area located in the left superior temporal gyrus (BA 22), posterior to Heschl's gyrus (primary auditory cortex). The PT is part of the notoriously known Wernicke's area and in neurotypical population shows leftward asymmetry in multiple aspects. Regarding the gross anatomy it was shown to have a greater amount of gray matter (e.g. Watkins et al. 2001) and greater surface area (Chance et al. 2006) in the left hemisphere than in the right hemisphere. With respect to the minicolumnar structure, it was repeatedly showed that PT has wider minicolumnar spacing (Buxhoeveden et al. 2001, Chance et al. 2006) and also consists of a greater number of minicolumns (Chance et al. 2006) than its right hemispheric analogue. Additionally, these differences were not found in Heschl's gyrus, therefore the wider minicolumn spacing in PT was interpreted as contributing to the right ear advantage - the left hemisphere superiority in processing of sounds with faster temporal change (e.g. Hickok 2009, Sætrevik & Specht 2012) and reversely, narrower spacing in the right hemisphere was interpreted as contributing to better spectral processing (Chance et al. 2006, Chance 2014). We will explain how the minicolumnar is spacing is thought to account for these specializations bellow.

Similar microanatomical results were found in fusiform gyrus, where again left hemisphere was observed to have wider minicolumn spacing and wider minicolumns, and right hemisphere was observed to have narrower minicolumn spacing and thinner minicolumns (Chance et al. 2013). When we compare the results from fusiform gyrus and PT, we can clearly see that there is a similar trend in the minicolumnar architecture of the hemispheres with left hemisphere having larger minicolumn spacing than the right hemisphere, yet the hemispheric specializations are found in contralateral hemispheres, since face recognition is processed predominantly in the right hemispheric fusiform gyrus but sound processing usually shows leftward advantage (Chance et al. 2013).

Differences in minicolumnar spacing are thought to be one of the main determinants of intercolumnar connectivity in hemispheres. It was proposed that in the left hemisphere, where generally greater spacing is observed, the connectivity is lower than in the right hemisphere with more closely packed minicolumns (e.g. Chance et al. 2013, Chance 2014, Jung-Beeman 2005). There is some anatomical evidence for this notion based on measurements of dendritic branching. Scheibel (Scheibel et al. 1985, Scheibel 1988) compared dendritic branching in multiple areas of the brain and found that neurons in the left frontal operculum (part of the Broca's area – left inferior frontal gyrus) have a larger proportion of higher order dendritic branches than neurons in the right frontal operculum, which mainly have lower order branches. Similar results were reported in orofacial motor area. Additionally, total dendritic length was greater in left frontal operculum and right motor orofacial area. Of course, total dendritic length alone doesn't necessarily tell us so much about connectivity in an area, due to factors such as minicolumnar spacing. However, if there would be similar total dendritic length in both hemispheres of some area, the connectivity would, of course, dependent on the spacing since the hemisphere with larger minicolumn spacing would have to grow longer dendrites to reach surrounding minicolumns and therefore would have a smaller total number of dendrites.

Indication that findings listed above are relevant for the language research comes from the fact that in chimpanzees, fusiform gyrus (Chance et al. 2013), and PT in both chimpanzees and rhesus monkeys (Buxhoeveden et al. 2001) doesn't show these asymmetries, with both hemispheres showing properties typical for humans' right hemisphere. Lack of microstructural asymmetry in chimpanzees' fusiform gyrus was explained as at least partially responsible for chimpanzees' enhanced face recognition ability (Chance et al. 2013). Additionally, unlike for minicolumnar asymmetry there is evidence that, for

instance, PT (Gannon et al. 1998) or BA 44 (Toga & Thompson 2003) in the left inferior frontal gyrus show leftward size asymmetry in great apes. These results therefore indicate that although there are some gross similarities between the humans and chimpanzees, minicolumns may represent the main innovation allowing the complex language and therefore are important for our investigation.

In this very quick outline of the minicolumn theory, we have seen that there are differences in the hemispheric specialization for certain types of processing and there are different underlying minicolumnar architectures in the hemispheres which are used to account for these processing type differences. Now we will explain how these microanatomical differences are thought to facilitate different processing types and figurative language processing.

General and maybe somewhat oversimplified view is that the left hemisphere facilitates processing that can be described as being more discrete, rigid, or more analytical. According to Jung-Beeman's (2005) generalization of minicolumn architecture, in the left hemisphere there is a greater distance between the minicolumns, dendritic fields are less overlapping and this relative sparsity of connections facilitates more discrete processing. An example of more discrete processing is, for instance, speech recognition or recognition of any sound with fast temporal change in general which are facilitated by the above mentioned PT in the left hemisphere.

Right hemisphere on the other hand is generally thought to be involved in more holistic processing, in creativity and more diffuse thinking. Minicolumns that are less distant in the right hemisphere have more overlapping dendritic fields and therefore this architecture facilitates this, using Jung-Beeman's (2005) vocabulary, coarser coding. An example of coarse coding is, for instance, face recognition facilitated by the above mentioned fusiform gyrus in the right hemisphere or better spectral sound processing in the right hemisphere.

These proposals are directly followed by the notion that (novel) figurative language processing that supposedly takes more diffuse and diverse thinking is facilitated by the right hemisphere. This is due to the fact that in cases such as novel metaphor processing two concepts that are semantically distant are more easily "connected" in the right hemisphere due to increased connectivity in comparison to the left hemisphere. This is supported, for instance, by the study of Faust & Mashal (2007) who found, using divided visual field paradigm, that the right hemisphere shows advantage for novel metaphors, but

not for conventional metaphors or literal word pairs. Other study of Kiefer et al. (1998), observed N400 event-related potential during indirect priming[7] and showed that the effect of indirect priming on this potential is present only in the right, but not in the left hemisphere.

Another evidence for this proposal comes from the studies of neuropil variations in psychopathological conditions that show aberrations of minicolumnar architecture and semantic processing, such as schizophrenia spectrum disorders. Schizophrenics, in contrast to neurotypicals, show indirect priming even with 0 ms inter-stimulus interval (Spitzer et al. 1993), they also tend to interpret word pairs that neurotypicals consider meaningless more often as meaningful (Zeev-Wolf et al. 2014), and it was shown that cognitive disorganization (assessed with schizotypy questionnaire) correlates with increased susceptibility to indirect priming (Johnston, Rossell & Gleeson 2008). Post-mortem studies of patients with schizophrenia patients usually show increased minicolumn density which is described as being caused mainly by the reduction of neuropil space between the minicolumns (Selemon & Goldman-Rakic 1999). Chance, Walker & Crow (2005) reported that this condition is accompanied with reduced density of GABAergic inhibitory interneurons in PT as well as in cingulate and prefrontal cortex, and authors hypothesized that reduced inhibition, supposedly resulting from such aberrations, can account for the symptoms of schizophrenia. Possible implication about the neurotypical population would be that if the neuropil reduction which is associated with increased intercolumnar density would be accompanied by decreased number of inhibitory interneurons, then the right hemisphere would show less inhibition than the left hemisphere and hence facilitate also the figurative language processing. This interpretation is in accord with the standard presumptions about the right hemisphere and the above mentioned evidence.

Another interesting phenomenon to consider was reported by Casanova (2007) who claims that reduced minicolumn spacing causes neuron cell bodies to be smaller. Since longer white matter fibers require larger cell bodies, this implies overall more localistic connectivity of a network with decreased minicolumn density (Casanova 2007). Some support for this notion comes from diffusion tensor imaging (DTI) studies. For instance, arcuate fasciculus (AF), a white matter tract connecting directly posterior temporal areas

[7] Example of an indirect priming stimulus is, for instance, the prime-target pair *sweet-lemon*. The prime and the target are here connected with the mediating word *sour* (sweet-sour-lemon). Indirect priming is a standard method to test free spreading of activation and is thought to be indicative of more diffuse processing than the standard, direct priming.

with inferior and superior frontal areas and indirectly the same areas with so-called Geschwind's territory in the inferior parietal lobe (Catani et al. 2005), consistently shows significantly higher radial diffusivity (RD) and lower fractional anisotropy (FA) in the right hemisphere than in the left hemisphere of neurotypical individuals (Upadhyay et al. 2007). However, in schizophrenics, increased RD and decreased FA were observed also (but not only) in left hemispheric AF (Leroux, Delcroix & Dollfus 2014). Such values are usually interpreted as evidence of lower number of axons in the tract (lower FA), and decreased myelination (higher RD). Since axons observed in large interlobar tracts such as AF are necessarily long, these tracts are not identified as consistently in the right hemisphere of neurotypical individuals, because this hemisphere's neurons should have smaller bodies, unable to metabolically support these structures. Evidence from schizophrenic individuals supports this notion since it shows anatomical aberrances we would expect to accompany extremely diffuse thinking.

To summarize the various evidence listed above we can assume that a neural network such as the one in the right hemisphere consists of more densely packed neurons with higher number of short dendrites and therefore shows more local than global connectivity and possibly also shows less inhibition. Of course, for the left hemisphere the situation is reversed. This is in accord with intuitivistic expectations of what the proprieties of the left and the right hemisphere should be like. However, when we look more closely at the evidence from research on autism, we will see that there is other variable that can cause a network with densely packed minicolumns to facilitate highly rigid processing.

7.3 Autism and schizophrenia

Autism and schizophrenia spectrum disorders are important for our inquiry for at least two reasons. First, and maybe the more important, is that both these disorders show abnormalities in metaphor comprehension and therefore we can draw some valuable inferences about how the neurotypical brain works as we briefly tried to do in the previous section. Secondly, a good model of figurative language processing should, ideally, be modifiable to account for these disorders and therefore a comprehensive neural theory trying to explain figurative language processing should predict the difference in the language processing when neural substrate changes. In this section we will focus mainly on the latter.

Behaviorally, autism and schizophrenia are quite the opposite when considering figurative language. Autistic individuals, in general, have difficulty to understand novel metaphors, as well as jokes, irony and other kinds of what we can label as creative and ambiguous uses of language (Vulchanova et al. 2015). Difficulties with metaphor comprehension are also reflected in electrophysiology by autists having more negative N400 for both novel and conventional metaphors (Gold, Faust & Goldstein 2010). Faust & Kennett (2014), in their network theory-inspired account, describe a semantic network of Asperger's and autistic individuals as highly rigid, meaning that it is a network with low connectivity (low mean vertex to node ratio) which causes larger average distance between concepts, a kind of extreme version of neurotypical left hemispheric network as it was described in previous section (although notice that these authors describe semantic network and not neural architecture). This is in accord with the usual notion of right hemisphere dysfunction in autism spectrum disorders which is reflected for instance in absence of right hemisphere advantage in novel metaphor comprehension (Gold & Faust 2012).

Schizophrenia spectrum disorders often show impairment that is, on the behavioral level, describable as opposite to the one we can observe in autism spectrum disorders – in schizophrenia we observe, what we could call overly loose interpretations or overinterpretations. Utterances which neurotypical individuals consider meaningless can be meaningful to schizophrenic individuals. For instance in the study of Zeev-Wolf et al. (2014) schizophrenic individuals and neurotypical controls had to indicate whether the presented (literal, metaphorical or unrelated) word pairs were meaningful or not. Schizophrenics rated approximately 50% of meaningless (unrelated) pairs as meaningful, whereas controls rated only 10% of these as meaningful. Semantic network that would facilitate such loose processing as observed in schizophrenia is in terms of Faust & Kenett (2014) describable as chaotic. In chaotic network there is abundance of connections and therefore there is very short average distance between the concepts.

Interesting view in this research line (Kenett & Faust 2014; Kenett, Anaki & Faust 2014) is proposal of a continuum where schizophrenia and autism represent opposing borderline cases. Other, less severe cases are represented for instance by schizotypal personality disorder or Asperger's syndrome and non-pathological cases are described with the notion of creativity – more creative people are found on more "schizophrenic side" of the spectrum and less creative people are found on more "autistic side" of the spectrum relative to some idealized neurotypical equilibrium. Kenett, Anaki & Faust (2014) tested

this proposal – they used free association task where participants had to list as many associations as they could come up with for a target word within one minute. Based on these responses, the researchers constructed graphs – association correlation networks and compared them between the groups of less and more creative people (creativity levels were assessed independently). Indeed, they found a significant difference between the networks, with more creative group having shorter average distance between nodes (interpreted as having shorter distance between concepts), higher clustering coefficient (higher probability that two nodes/concepts will be neighbors) and less modularity than the less creative group (Kenett, Anaki & Faust 2014). These results are consistent with the expectations of more chaotic semantic network in more creative individuals and more rigid semantic network in less creative individuals. Although these differences were significant, they were rather small, but we need to keep in mind that this study was conducted exclusively on neurotypical participants and needs to be conducted on autistic and schizophrenia spectrum individuals to prove this promising proposal truly valid.

Based on these observations of autistic processing being overly literal or rigid, and schizophrenic processing being overly loose or chaotic, straightforward hypothesis about minicolumnar architectures underlying these impairments would postulate that in autism we will observe overly sparse minicolumnar spacing and in schizophrenia this will be overly narrow. However, this is only partially true since in both disorders, we observe somewhat similar trend of microstructural aberrances.

In schizophrenia, as already mentioned, we often observe more densely packed minicolumns. For instance Chance et al. (2008) observed reduced minicolumn spacing in PT of both hemispheres in males (although opposite was observed in females). In the left hemisphere the size of PT's surface area was smaller and correlated with minicolumn spacing. This is indicative of repeatedly observed decreased functional and structural hemispheric asymmetry in schizophrenia (e.g. Oertel et al. 2010). This effect of reduced asymmetry can be very strong and for instance Hasan et al. (2011) found reversed volume asymmetry of PT in first-episode schizophrenics. Notably, Oertel et al. (2010) found that decreased functional asymmetry correlates with increased severity of schizophrenia symptoms. Studies of schizophrenic patients usually report decreased gray level volume (e.g. Chance et al. 2008, Casanova et al. 2008, Oertel et al. 2010) which is indicative of decreased size of neurons that is thought to be connected to reduced minicolumn spacing as noted above (but might also indicate decrease of total number of neurons). Decreased

asymmetry in schizophrenia is also often reported as a consequence of mainly left hemisphere impairment (e.g. Hasan et al. 2010) which is consistent the coarse coding theory since it renders left hemisphere to have more typically right hemispheric properties.

Brains of autistic individuals, contrary to intuitions built by the coarse coding theory don't show increased, but decreased minicolumn spacing which is finding similar to what was found in schizophrenia patients. For example, a two-case study of Buxhoeveden et al. (2006) showed that frontal cortex of autistic patients contains more narrow minicolumns than it does in neurotypical individuals, and this was most pronounced in its dorsal and orbital parts. This is important, since if we take into account that cell bodies are smaller when minicolumns are more densely packed, yet *„increased brain size appears to be the most consistent morphometric observation reported in autism*"(Casanova 2007), we can see that this is possibly due to increased total number of minicolumns (Casanova 2007). Therefore in autism we observe hyperconnective network just as in schizophrenia but with increased number of processing units. This view is even strengthened by a recent study of Hutsler & Zhang (2010) observing dendritic spite anomalies in the brains of autistic spectrum disorder individuals. The authors reported increased density of dendritic spines in areas of frontal, temporal and parietal cortex in autistic spectrum individuals compared to neurotypical individuals. The authors explain this observation in terms of a proposal claiming that autism spectrum disorder's symptoms are caused by reduced long-distance connectivity and increased short-distance connectivity (Hutsler & Zhang 2010), sometimes termed as the increased local connectivity hypothesis of autism. Evidence for this proposal can also be seen in some DTI studies, for instance Fletcher et al. (2010) showed that AF in high-functioning autistic adolescents is disrupted bilaterally and review of Hoppenbrouwers, Vandermosten & Boets (2014) reports white matter disruption in autism in general and thus supports this view.

Some authors argue against the increased local connectivity hypothesis of autism. For instance, it was found that short-distance white matter tracts which would be expected to be intact or enhanced in a condition with local hyperconnectivity are impaired in autism as well (Shukla et al. 2011; Hoppenbrouwers, Vandermosten & Boets 2014). The problem with this notion is that it considers short-distance white matter connectivity equal to local hyperconnectivity. As mentioned earlier, original arguments for the notion of local hyperconnectivity are based on increased number of short distance dendrites, dendritic spines and decreased intercolumnar spacing. These factors alone may be significant

enough to overcompensate for decreased structural consistency of axons. What needs to be considered as well is the fact that DTI measures water diffusion in voxels, not axons per se. If there is no principal direction of axons in a voxel due to a chaotic organization then we would expect observe the same abnormalities in short-distance white matter tracts using this technique. Based on these methodological doubts, we will still assume that the increased local connectivity hypothesis of autism is valid.

7.4 Overfitting and underfitting proposal

Based on the above mentioned evidence, hypotheses, and theories, it is possible to formulate a simple description of autism and schizophrenia that will later help us to explain aberrant semantic processing observed in these conditions, using tools provided by an integrative account presented in the next chapter.

As we could see there is a lot of what schizophrenia and autism have (or are expected to have) in common at microanatomical level – reduced intercolumnar spacing, decreased neuropil volume, smaller cell body size and resulting local hyperconnectivity and global underconnectivity. What seems to be the major difference in this respect is the total number of neurons which is an approximation based mainly on observations of different brain sizes. As mentioned earlier, brains of autistic individuals are consistently observed to have larger size than brains of neurotypicals (Casanova 2007). Given this fact and the decreased intercolumnar spacing, we can easily say that autists' brains are composed of a higher number of functional units than the brains of neurotypicals. On the other hand in schizophrenia the brain is often smaller (e.g. Selemon & Goldman-Rakic 1999, Steen et al. 2006) and therefore is composed of the same or lower number of functional units than neurotypical brain.

We hypothesize that autism and schizophrenia can be described in terms of the *overfitting* and *underfitting*, terms usually used in machine learning and statistics. Overfitting describes a state when a model is too complex and therefore excessively precise, describing not just the trend in the data, but also the noise. It is quite obvious that autism can be described in terms of overfitting. Autists construct a model of world that is too precise and therefore any aberration from this precise overlearned model is considered an error. In natural language processing this is manifested by too precise or we can say highly constraining model of words' interpretations and therefore a decreased variability of possible words' meanings. If an actual interpretation of a word is different from an

expected interpretation it is very likely that it will be considered erroneous. This is precisely the case of novel metaphors where the deviation from expected literal interpretation is even larger than in loose literal interpretations. Therefore when an autistic individual encounters a novel metaphorical utterance he will either find an interpretation that is overly literal but in accord with this overfitting model or he will refuse an utterance as meaningless if it is not possible to find an interpretation consistent with the model. Since this model is so restrictive in comparison to the model of neurotypicals, a significantly lower quantity of data can fit its constraints. This can account for instance for well-described increased discrimination but impaired generalization. Phenomenologically, this is reflected in autists' feeling of being overwhelmed by sensory information (e.g. Pellicano & Burr 2012). This is no surprise given that if a model is too restrictive it cannot account for much variation in the data and therefore a large proportion of the data will be considered "new". Some indicative evidence that this proposal may be valid comes from the field of artificial neuronal networks, where one of the methods to avoid overfitting is so-called *dropout* method (Srivastava et al. 2014). When this method is used, in every learning epoch a part of network's neurons along with all their connections is temporarily removed, based on probabilistic selection. Important for us is that overfitting can arise from overly high number of functional units which naturally adds to complexity. By reducing this complexity, overfitting decreases. Analogy with abundant amount of neurons observed in autism is obvious and supports the view of autism as overfitting. This analogy also indicates why it is so that although in autism we observe many microanatomical features similar to schizophrenia, individuals with these disorders behave differently to the point that these disorders are considered opposite.

We should mention that we are not the first to think about overfitting in context of autism and the same basic idea can be found in the letter of Bakouie, Zendehrouh & Gharibzadeh (2009) and the comment of de Cruys et al. (2013). These authors came to the same conclusion although they only described this phenomenon at behavioral, not microanatomical level.

Schizophrenia is describable as so-called *underfitting*. Underfitting refers to a state when a model is too simple and therefore it is unable to properly describe more complex data. Often this is used to describe a model that is erroneous due to this simplicity, but we will use it in a slightly different sense as described for instance by Alst et al. (2010) - underfitting "*allow[s] for much more behaviour without strong support for it"*(Alst et al.

2010). It is therefore such a model that is excessively simple, hence less constraining and consequently fitting more data. This notion is analogical to schizophrenics' behaviour and explains why they find meaningful many metaphorical word pairs neurotypicals find meaningless (Zeev-Wolf et al. 2014) and also, for instance, schizophrenics' general failure to find mistakes in semantic and discourse level stimuli (Stephane et al. 2007). As already described, this is probably facilitated by hyperconnective network, but of course with less functional units than in autism.

Because there is little discrimination, but not an abundant number of functional units, schizophrenia, but not autism can be *directly* described in terms of previously mentioned Kennett's & Faust's (2014) network model as chaotic. In the case of autism, rigid network can be considered a good model to describe it at behavioral and possibly phenomenological, but certainly not at implementational level, since the neural network is chaotic in this sense as well.

We will return to the problem of overfitting and underfitting in autism and schizophrenia in chapter 8.3.2 where the above notions of autism as overfitting and schizophrenia as underfitting will be integrated within the theoretical framework described in the next chapter.

8 Theoretical integration

In this chapter, we will try to integrate findings and observations mentioned in previous chapters using mainly the relevance theory and the career of metaphor theory but partially also the class-inclusion theory. As we will show, these views are not particularly difficult to integrate. The integration we will propose has some emergent qualities, mainly it can help us to fuzzify the dichotomy of categorization and comparison and it can account, although in a very basic way, for metaphor processing in autism and schizophrenia. Sometimes, we will use very simple formalisms to present our ideas more clearly, but the theory doesn't stand on these.

We will assume the same type of knowledge representation as the career of metaphor theory. Therefore, similarly to this theory, the basic assumption we will work with is that it is possible to represent human conceptual system in terms of a propositional network and that the rules of structure mapping theory apply as well. Concepts can be defined as subsets of conceptual system and hence they are of propositional nature as well. Motivation to use this knowledge representation is, as we mentioned repeatedly, that to our knowledge the career of metaphor theory is the only viable psycholinguistic theory of metaphor that is built on a well-developed ontology and so far this ontology seems to be valid. The conceptual metaphor theory also proposes an appealing ontological account on its own, but we won't consider this theory for reasons mentioned in chapter 4. To a reader interested in accounts that integrate relevance theory and conceptual metaphor theory, we recommend the works of Tendahl & Gibbs (2008) and Wilson (2011).

What is more important in this theoretical integration than concepts themselves, are interpretations. Interpretation here is a pairing of a word form (further just 'word') and a conceptual structure. We will talk about *possible interpretations* and *actual interpretations*. An actual interpretation is a correct interpretation of a certain word in a certain context. Possible interpretation will be defined as an interpretation that has probability of being an actual interpretation higher than zero, but is not necessarily an actual interpretation.

8.1 Lexical loosening and propositional networks

As already mentioned in chapter 5, the relevance theory proposes that even the most ordinary literal words can encode a wide variety of interpretations. Consider these examples from Wilson (2011):

(16) The play starts at 7.00.

(17) Jane's hair is straight.

It is quite clear that saying that something starts at 7.00 is usually a mere approximation and also that Jane's hair is obviously not straight in the geometrical sense, and therefore this utterance's meaning can be considered approximate as well. These cases are, in terms of the relevance theory, instances of so-called lexical loosening. As mentioned in chapter 5, metaphor in the relevance theory is considered to be an instance of lexical loosening as well. We will demonstrate that this applies to metaphors on the previously mentioned word *bank.* This word can be initially interpreted as some bank institution in general, but by lexical loosening, it can also be interpreted as any financial institution that provides services largely similar to banks' services. Now consider the sentence (18):

(18) My father is a bank.

This sentence can mean that the speaker's father has a lot of money and it is usual that he gives it to the speaker. To interpret this sentence in this way, it is needed to loose the interpretation even more than in the above mentioned example of a non-bank institution, since it communicates only a very small amount of properties in comparison to other loose interpretations or to the original interpretation of this word. Based on the structure mapping theory, we can assume that all these interpretations can be represented as propositional structures. Important for us is how the difference between degrees of lexical loosening is reflected by propositional structure of an interpretation. As we can see on the loose interpretations of the word *bank*, the looser the interpretation, the fewer elements it has. This is in accord with the structure mapping theory's classification of comparison types – if the base and the target would share all their features, we would be talking about identity, but since this is not the case and only subsets of the original base concept are matching the target, with decreasing number of matching elements the definition of comparison type shifts from literal similarity to nonliteral types of similarity, such as metaphor. The same rule applies to categorization in general, consider for instance the literal category *ANIMAL* and compare it to any member of this category. Every member of this category has all the properties of the concept *ANIMAL* plus species-specific properties.

With this in mind we can also describe lexical narrowing, which is opposite in direction to lexical loosening, since it is a type of interpretation shift that contains more elements than

the original concept. For the word *bank*, an example of lexical narrowing is the use of the word to refer to some specific bank.

It should be noted that there are cases of lexical loosening that cannot be described as decreasing in number of properties, such as the word *straight* in (17), number *7.00* in (16), or use of a word *green* to describe a group of things that are not necessarily green. This framework doesn't explain loosening of these, in a sense, perceptually almost atomic concepts, but rather basic level concepts made of a high number of properties, such as the word *bank* in (18). This is because it is hard to imagine how would be loosening of these atomic concepts described in terms of decreasing number of properties/nodes. Further in the text, we will sometimes still use these atomic examples, but only in contexts where this notion is irrelevant.

Now we can see that the relevance theory's notion of lexical loosening and narrowing is consistent with propositional network knowledge representation without almost any problems. However, for the sake of further analysis, there is one aspect of the relevance theory that needs to be adjusted. The problem is based on the question of what determines which interpretation represents the reference point for lexical loosening and narrowing. Relevance theory assumes that in the case of the word *bank*, it is the general notion of a bank, in (17) it is the concept of straightness in its strict geometrical sense. In other words, it assumes that there is some *linguistically encoded concept* which is encountered first and serves as a cue to interpretation, which is then achieved through the process of lexical loosening and narrowing. However, it seems that this linguistically encoded concept is in this theory conceived as rather rigid since all metaphors, but also other conventional words' interpretations, such as the word *Xerox* which is conventionally used to refer to a printer, are considered here loose interpretations and hence derived from some linguistically encoded concept. Consider again the word *straight* in (17). It is not very plausible to think that people (who are not geometricians) would during the processing of (17) think first about the geometrical concept of straightness and then derive the loosened meaning. Rather, people will interpret it immediately with a concept that is described by the relevance theory as loose, because it is used more often in such a context and therefore it is more likely to be an interpretation of the word *straight*. If this interpretation fails, loosening or narrowing will occur. This argument is basically a paraphrase of the relevance-theoretic comprehension procedure (Wilson & Sperber 2006), which also postulates *"a path of least effort"* (Wilson & Sperber 2006), with the slight difference that

it doesn't assume the existence of some linguistically encoded concept shared across all contexts to be necessarily activated.

Since from all possible interpretations, the individual will always select predominantly the one that is most probable, such an interpretation should be considered a reference or a starting point of lexical loosening and narrowing, although it may differ from the textbook definition. We will call this type of interpretation *default interpretation* and define it in probabilistic terms as a conceptual structure *K* that will most likely be encoded by a word *S* in a context *C*. This implies that with different context, this probability may change. For instance, the word *shark*, if it is conventionally used metaphorically as well as literally, may have different default interpretation when a person watches a documentary movie about these animals and when he is listening about his overly aggressive and ambitious boss. This, however, doesn't mean that when a person is in a context that primes the metaphorical interpretation, the literal interpretation will be completely neglected. If both interpretations are used frequently, although in different contexts, an interpretation that is currently out of context may still serve as a strong possible interpretation. This could be described as trade-off between overall and context-dependent use of an interpretation.

Importantly, we don't want to claim here that there is necessarily only one default interpretation in a single context. For instance homonymous or polysemous words and words that are used conventionally as metaphors (remember from chapter 5 - an interpretation to be perceived as metaphorical must have recognizable literal meaning) may have multiple default interpretations if the context is not clear or is ambiguous or both these interpretation are conventionally used in the same context – generally, in cases where context doesn't provide sufficient resolution.

Obviously it is not always the default interpretation that is an actual interpretation of a certain word, and in such situations a loose or narrow interpretation must be found in order to understand a word. First, we need to set a general rule that every *possible interpretation* is describable as probability that some conceptual structure *K* will be encoded by a word *S* uttered in a context *C*, or $I_n = P[(K|S)|C)]$, such that $I_n > 0$. Note that this is just a generalization of what was said earlier about the default interpretation. From this follows that the default interpretation I_α is describable as $I_\alpha = maxP[(K|S)|C)]$. When a default interpretation fails to be an actual interpretation of a word, the second most probable

interpretation is selected[8] and so on, until an actual interpretation is found or the comprehension procedure becomes too effortful compared to the expectation of relevance (Wilson & Sperber 2006). The general idea is illustrated on figure 1 with probability density function.

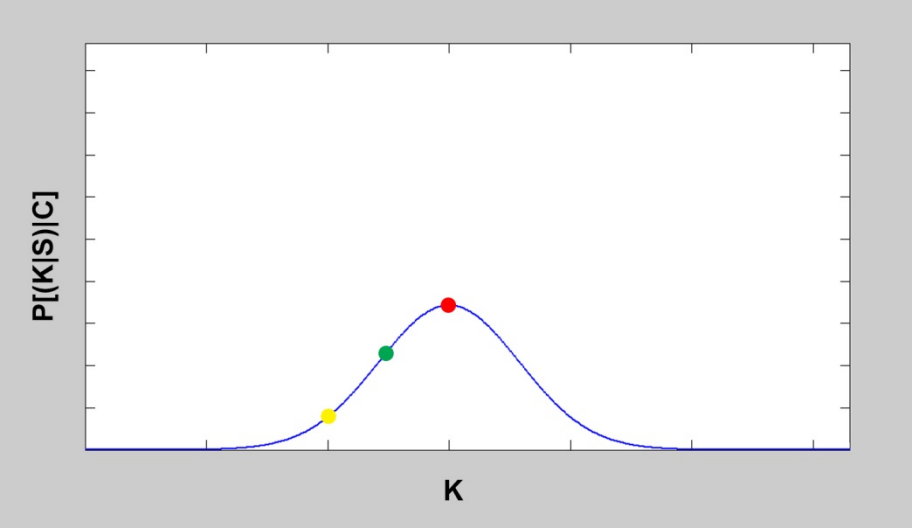

Figure 1 Hypothetical probability distribution of possible interpretations of a certain word with single default interpretation in a certain context. Red dot – default interpretation; green dot – loose literal interpretation; yellow dot – (loose) figurative interpretation.

On the x-axis conceptual structures are ordered from most loose to most narrow. Loose interpretations, although they can be used to describe larger quantity of data than their default interpretations, are smaller in number of properties and therefore on the x-axis interpretations are ordered from smallest to largest conceptual structures. On the y-axis we see the probability of every conceptual structure in our window being an actual interpretation of a certain word in a certain context. Please note that the depiction on figure 1 is only illustrative – there can obviously be a high number of conceptual structures that are one property smaller than some other conceptual structure, but in this example, for simplicity, we assume that there is only one conceptual structure of certain size for a word.

[8] For simplicity, we will usually expect that possible interpretations are selected from the space of loose interpretations and therefore the second most probable interpretation is the first most probable possible loose interpretation. In the current state, this theory is unable to specifically describe under what circumstances a hearer decides to search an actual interpretation in the space of loose or narrow interpretations.

Let's assume that word to be interpreted is the above mentioned word *bank* and that this word doesn't conventionally encode any conceptual structures other than the basic one which is interpreted as some bank institution. This conceptual structure *BANK* then serves as the default interpretation of the word *bank*, since it is most probable interpretation of this word (in some context C or all contexts, but that is not important now). On figure 1, the default interpretation is indicated by the red dot. Loose interpretations are indicated by green and yellow color, and in this case the more loose the interpretation, the less probable it is to be an actual interpretation of the word *bank*. Interpretation indicated by the green dot represents, in this hypothetical case, the interpretation that describes also financial nonbank institutions and yellow dot represents the metaphorical case from (18).

However, this is just an example and this probability distribution can be different for different words. For instance for the word *shark*, we could expect a bimodal distribution since both metaphorical as well as literal interpretation are conventional. Since this distribution can change with learning, there is no reason to propose that metaphors are necessarily *ad hoc* concepts, unless they are novel. As mentioned above, default interpretation of the word *shark* in the sentence *My boss is a shark* uttered in an appropriate context is a metaphorical interpretation. With this notion of learning as change of probability distribution of possible interpretations, the concept of conventional metaphors or metaphorical category learning is integrated with the dynamic view of the relevance theory.

What we did not define yet is context. When talking about mind and brain, every brain state other than the neural activation pattern of an interpretation itself, can be understood as a context if it has modulatory influence on the interpretation selection. Based on this, we can define noncontextual interpretation of a word as an interpretation of a word activated when there is no modulatory influence on the concept selection other than the word form alone. Context is then any brain state that has an influence on the concept selection other than the brain state caused by the word form alone. An universal example of noncontextual interpretation is difficult to provide since there may always be a number of inner or outer reasons that can shift an interpretation, especially if a word has multiple default interpretations. However, we can roughly imagine it as some single noun utterance without known illocutionary intention from unknown source and not of particular emotional value. Epistemologically this notion is probably not valid and for now remains a theoretical

possibility. The notion of noncontextual interpretation $C_{Ø}$ is motivated by the above mentioned general rule from which follows that if $C = C_{Ø}$, then $P[(K|S)|C)] = P(K|S)$.

8.2 Categorization and comparison revisited

The notions above don't provide us with tools sufficient to disentangle the problem of categorization and comparison in metaphor comprehension. To do this we have to look at the study of Tenenbaum & Xu (2000) which described word category learning as Bayesian inference. In this study, participants were learning word categories from visual stimuli (photographs), which were provided with artificial category names. Participants' implicit task in this learning phase was to infer the properties of a category, based on co-occurrence of an artificial category name with some visual stimuli (for instance three photos of different dogs). After this test phase, among other tasks, participants were provided with random pairs of stimuli from the learning phase and asked to rate how similar the depicted objects are based on criterions they used in the learning phase. Using hierarchical clustering algorithm authors showed that more distinctive clusters of objects are more likely to have distinguishing names (category names). In other words the greater the vertical distance (distinctiveness or difference) of a parent concept from its child concept, the higher the probability that they will have distinguishing names (Tenenbaum & Xu 2000). We can demonstrate this on the example of categories *MAMMAL* and *CAT*. There is a large difference between the number of properties encoded by these concepts and the *MAMMAL* is obviously a hypernym of the *CAT*, with *CAT* having a large number of specific properties that *MAMMAL* doesn't have. The properties of *MAMMAL* therefore represent only a relatively small subset of *CAT*'s properties. Now imagine that a person's definition of a cat contains properties of having big eyes, being sneaky and having a fur. If we would show to this person an Egyptian furless cat, he would either define it as similar to a cat, or change his definition of cat, so it does not contain a property of having a fur. This new category would be hierarchically higher, compared to the original category of furry cats, since it would be one property smaller. It is unlikely that the person would keep both furry *CAT* category and not-necessarily-furry *CAT** category as two distinguished categories with different names although by the definition *CAT** is, of course, a hypernym of *CAT*. The same way, a person would not create a new distinguishing name to a narrowed, one property richer, category *CAT*** which describes all black cats. *CAT*, *CAT** and *CAT*** would standardly still be referred to by the word *cat*, although, of course, it is possible to distinguish these in other ways when needed. However, a specific well-known

cat named *Tiger*, will have a specific name for a reason – there is a lot in *TIGER* that is not in *CAT*, with all Tiger's specific behaviors and the specific way he looks.

It could be argued that this study merely tells us that category labeling with specific names depends on the degree of difference between vertically aligned concepts. However, we can hypothesize that the greater the difference between two vertically aligned concepts, the higher the probability that these will be recognized as having distinguishable hierarchical relationship, in other words, that the perception of hierarchical relationships will depend on degree of this difference. We will further develop and test this hypothesis in context of metaphors.

Inspired by the study of Tenenbaum & Xu (2000), we can look at the problem of categorization and comparison from two perspectives, first and weaker being the difference between original literal and loose metaphorical base concept, and second and stronger being the difference between base and target concept.

The class-inclusion theory holds that an abstract metaphorical category is superordinate to both the target being described as well as the original literal concept from which the abstract metaphorical category was derived. If this is true, we can say that whether we talk about categorization or not, depends on how different is the abstract loose interpretation from its default interpretation. If this difference is small, we can expect that this metaphor will more likely be considered a comparison, but if the difference is large, it will be considered a categorization. Although this argument doesn't touch the relationship between the base and the target directly, it gives us some indication about the difference between the two processing types.

For stronger account, we have to ask, what is the difference between categorization and comparison considering both base and target? Categorization in framework of structure mapping theory can be described as such comparison type where target concept has all properties of the base concept and some more in addition (recall the previously mentioned example of category *ANIMAL* or *MAMMAL*). Therefore when we talk about figurative categorization, we are looking for some loose interpretation derived from an original literal base concept. The problem is that when we talk about comparison, we do the same thing since what is similar between two non-identical concepts is a subset of both the base and the target concept and therefore, again, we must find some loose interpretation of the base. Therefore we propose, inspired by the study of Tenenbaum & Xu (2000), that the main

difference between the categorization and comparison in metaphor comprehension is based on the difference between the number of properties encoded by the abstract metaphorical concept and the target. As the difference between the total number of properties that are encoded by the target and by the abstract metaphorical base concept increases (number of shared properties decreases), the metaphor is more likely to be interpreted as categorization (and vice versa for comparison). In other words, as the ratio between the number of properties encoded by the abstract metaphorical concept and the target increases, the difference increases as well and therefore this ratio is indicative of comparison or categorization. This ratio, of course, increases as the interpretation of a base concept is increasingly looser since the target remains the same size but the base becomes smaller. The general idea is illustrated in figure 2. On the base position, colored nodes[9] represent a loose metaphorical concept. The rest of the base nodes belongs to the original literal concept – the default concept. On the target position we can see the part of the target concept that is described by the base (colored nodes). The rest of target's non-colored nodes is indicative only of the difference between the number of properties described and not described by the base (target doesn't change during metaphor comprehension so this, of course, doesn't indicate lexical loosening). Notice that the orange target has more nodes described by the base (it is comparison) than the blue target (it is categorization).

It is also very likely that as the difference increases, the loose interpretation will be different enough to be considered hypernym of the original literal concept as well, because as the number of properties shared between the base and target decreases (as looser interpretations are selected), so does the number of properties between the literal and metaphorical interpretation of the base. Therefore if this is the case, the result is the same type of categorization as is proposed by the class-inclusion theory of metaphor (Glucksberg & Haught 2006). However, this does not always have to be the case. For instance, in novel metaphors it seem highly possible that it is often not entirely clear what aspects of the base concept should be projected on the target concept and hence it is also possible that in novel metaphors more properties is mapped, although with lower weight (which we could, in this particular case, describe as an expression of the degree of mapping's certainty), and therefore an interpretation is more similar to both target and

[9] Nodes are here used to refer to any kind of element described by the structure mapping theory – we operate only with the number of properties (elements), not with their specific qualities.

consequently such metaphor is more likely to be considered a comparison and therefore preferred in comparison form.

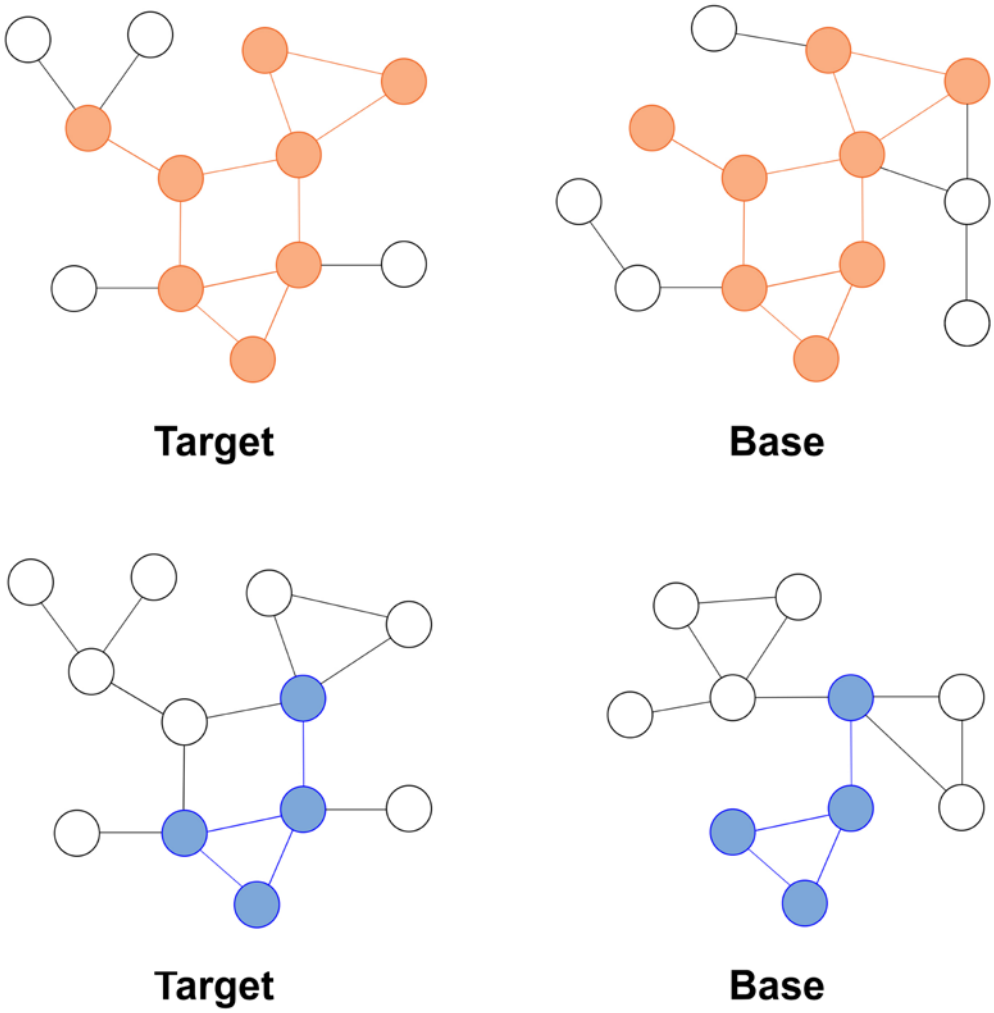

Figure 2 Illustration of the proposed difference between categorization and comparison in metaphor comprehension. The images above (orange) show comparison – as can be seen on the target, most of its structure is shared with the base. The images bellow (blue) depict categorization – significantly smaller proportion of its structure is described by the base in contrast with comparison.

In other words, we think that if this ambiguity is present, it is likely to cause that loose interpretation won't be different enough to be considered hypernym of the target. However, as a metaphor is used repeatedly, it becomes clear which of these originally mapped properties are valid and which are not, and this causes that an interpretation is

looser and therefore the metaphor is more likely considered a categorization. This, we think is why, for example, in vitro conventionalization (mentioned in chapter 3) induces switch of the sentence form preference from comparison to categorization form.

On the other hand, when a novel metaphor is very apt, meaning here that the properties from the base concept that are expected to be mapped are unambiguous and their number is sufficiently small, then even a novel metaphor can be readily conceived as categorization. Consider this rather rude metaphor that is sometimes used to describe that somebody is very tall:

(19) She is a giraffe.

This is extremely loose use of a word giraffe, since it contains only one property, being tall, yet it is easily understandable in categorization form even when it is novel.

8.3 Error

The concept of error is important for further analysis of the difference between novel and conventional metaphors as well as of the aberrant semantic processing which is observed in schizophrenia and autism. Error can be defined as the difference between expected state and actual state. Obviously, the larger is the difference, the larger is the error. Expected state in this theory is the most probable interpretation and therefore it is the default interpretation. If default interpretation is not an actual interpretation then error is the difference between the probability of the default interpretation and an actual interpretation. The larger is the error, the less likely it is that a possible interpretation will be an actual interpretation. In terms of the relevance theory, error could be conceived as an indicator of processing effort. If processing effort is larger than expected relevance of a stimulus (in our case a word or a sentence), then it is unlikely that such potential interpretations will be approached and hence an utterance won't be understood. Why should processing effort be higher for potential interpretations that are less probable to be actual interpretations? Since it is expected that more probable potential interpretations will always be processed prior to less probable potential interpretations, this increasing quantity of potential interpretations attended will obviously increase processing effort (although they may be, and probably are, other factors causing processing effort to increase).

8.3.1 Error in novel and conventional metaphors

With the notion of error, predictions about novel and conventional metaphors can be formulated more precisely.

An explanation of the comparison form preference in novel metaphors is quite straightforward – novel metaphors are preferred in comparison form because comparisons' conceptual structures are less distant (less different) from default interpretation than categorizations' conceptual structures and hence they are less erroneous than categorizations. People prefer less erroneous interpretations and therefore they prefer comparison form of novel metaphors. This is reflected in some informal observations achieved during the stimuli rating phase of the experiment reported in the next chapter. Raters in this experiment evaluated sentences in categorization form and they often mentioned that many of these sentences would be less "weird" or "odd", if they would be phrased as comparisons ("if there would be 'like'"). As we mentioned earlier, it might be the case that some novel metaphors are ambiguous, meaning that it is not clear which properties are meant to be mapped and therefore more properties, but probably with lower weight, are mapped. In such cases it may sometimes be even impossible to construct categorization. This, however, only explains why novel comparisons are preferred over novel categorizations and not how novel metaphors are processed in general. Since in conceptual system there is no predefined metaphorical interpretation prior to novel metaphor comprehension process that would guide this comprehension process, we think that, in general, novel categorization form sentences will facilitate construction of more distant interpretation than novel comparison form sentences. In other words, if hearer doesn't have any strong conventional metaphorical interpretation to converge to, he will predominantly rely on the sentence form.

Conventional metaphors are more difficult to explain within this theory. When conventional metaphors are presented in categorization form, a default interpretation, which is conventionally used as an actual interpretation of a word in figurative context, will be used. This default conventional figurative interpretation is assumed to be, at least in an usual case, distant enough be classified as categorization and hence conventional metaphors are preferred more often in categorization than in comparison form (Bowdle & Gentner 2005). However, it is questionable how conventional metaphors in comparison form will be interpreted. There are at least two possibilities: a) the same default figurative

interpretation will be used for both conventional categorization and conventional comparison form metaphors. In this case, the delay in processing time of comparison form sentences, compared to categorization form sentences (Bowdle & Gentner 1999, 2005) would be caused by an interference of the word *like*, since this would imply different relationship between the base and target than there actually is; b) different interpretations will be used for conventional comparison form sentences than for conventional categorizations. In this case, the word *like* would invite an interpretation less distant from the default *literal* interpretation a hence satisfy conditions of comparison relationship between base and target. This would be expected to take longer to process, since the most accessible interpretation, the default figurative interpretation, would not be used. The b) is probably more likely to be true since as for instance the study of Keysar et al. (2000), described in chapter 4, showed, even conventional metaphors rely heavily on their word forms (of course this is merely an indirect indication not an evidence).

Predictions listed in this section were tested using *free paraphrase method*. Design, procedure and results are listed in chapter 9.

8.3.2 Error in autism and schizophrenia

When considering autism and schizophrenia we propose that aberrant neural network architectures observed in these disorders, as they were described in chapter 7, lead to changes of possible interpretations' probability distribution and consequently to different error sizes which leads to aberrant semantic processing. We proposed that autism is describable as so-called overfitting, due to predictions being too precise or specific. In here presented model, overfitting is reflected by increased steepness of the probability distribution of potential interpretations. This causes the difference between every two potential interpretations' errors to be larger than in neurotypicals (compare the left part of figure 3 with figure 1). Consequently, looser and narrower interpretations are less accessible for autists than for neurotypicals which is manifested as overly rigid thinking. On the other side, in schizophrenia, as can be seen at the right side of the figure 3, the slope is more gradual than in neurotypicals which reflects the proposed underfitting. This causes the difference between every two potential interpretations' errors to be smaller than in neurotypicals. Consequently, looser and narrower interpretations are more accessible for schizophrenics than for neurotypicals, which causes for instance overly diffuse thinking and renders significant number of typically nonsensical sentences as meaningful.

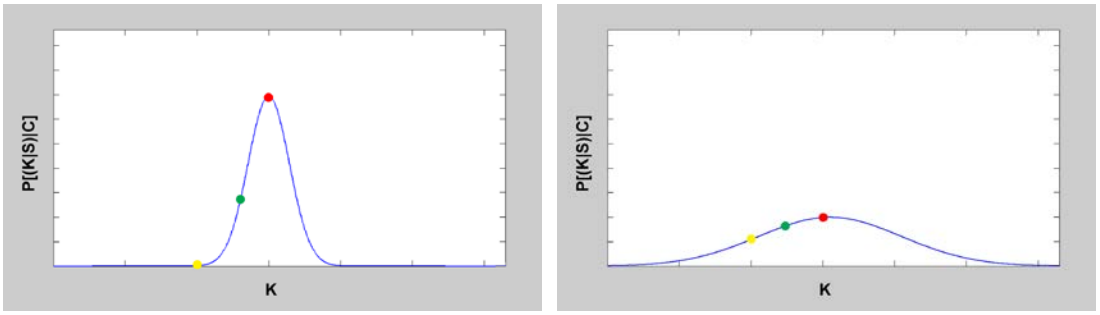

Figure 3 Hypothetical probability distribution of possible interpretations of a certain word with single default interpretation in a certain context in autism (left) and schizophrenia (right).

9 Experimental study

9.1 Motivation

This study was designed to test predictions proposed in the previous chapter (explicitly formulated in section 8.3.1). This proposal claims that novel metaphors will be interpreted differently when presented in comparison form and in categorization form, with comparison form metaphors' interpretations being less semantically distant from the original literal concept than categorization form metaphors' interpretations. This is based on the assumption that when there is no well learned interpretation of a word and therefore an *ad hoc* interpretation must be approached, the sentence form can influence how loose a selected interpretation will be, and therefore interpretations will often be different as a function of sentence form. Conventional metaphors will either be interpreted similarly in both forms due to conventional metaphorical concept being applicable to both comparison and categorization form sentences or differently as a function of the sentence form (again, with comparison form metaphors being less semantically distant from the original literal concept).

To test these predictions the free paraphrase method was used. When this method is used on metaphors, the task is to paraphrase the sentence in order to explain what it says about the target (basically, to provide an interpretation of the sentence). The target remains unchanged. For instance, if the sentence *Some ideas are diamonds* is used, possible paraphrase may be *Some ideas are rare and desirable* or S*ome ideas are brilliant and insightful* (Glucksberg & Haught 2006). To our knowledge, there are two studies which used this method that are of interest for this study. Both these studies observed the same variable, predication type of properties produced as paraphrases of metaphorical sentences. Two predication types, double and single predications, were recognized in these experiments. Double predication is such property that is applicable to both target and base. Example of a double predication is the above quoted paraphrase *some ideas are rare and desirable*. Both diamonds and ideas can be rare and desirable. Single predications are applicable only to target. Example of a single predication is the word *insightful* in the above quoted paraphrase *some ideas are brilliant and insightful*. Only ideas, not diamonds, can be insightful. Hasson and colleagues (Hasson et al. 2001 as cited in Glucksberg & Haught 2006) found that categorization form metaphors are more often paraphrased with single predications and almost never with double predications, but comparison form

metaphors are paraphrased with roughly the same number of single and double predications. The number of single predications in comparison form metaphors was lower than in categorization form metaphors. Authors interpreted this as evidence that categorization form sentences are "more metaphorical" than comparison form sentences. However, this study did not consider conventionality of metaphors therefore these results must be considered with caution. Also, we were not able to acquire the original study (Hasson, personal communication) therefore our report of its results must be taken with caution as well.

In another study, Bowdle & Gentner (2005) observed the same variable, but with a rather different setting. Participants had to provide a paraphrase as fast as possible (because this measurement was part of a study observing primarily reaction times). This study considered conventionality, but used the base-conventionality measurement which was criticized as possibly flawed (Thibodeau & Durgin 2011; see chapter 6). In any case, Bowdle & Gentner observed, unlike the previous study, no significant difference between categorization and comparison form. The only significant factor was conventionalization with conventional metaphors having higher number of single predications than novel metaphors. Authors interpreted, similarly to the previous study, double predications as indicators of comparison processing, but single predications as indicators of vertical (hierarchical) relationship, therefore supporting the view that novel metaphors are processed as comparisons, but conventional metaphors are processed as categorizations.

Our study examined the same phenomenon as the above mentioned studies with some corrections in the design. Unlike Hasson and colleagues (Hasson et al. 2001 as cited in Glucksberg & Haught 2006), this study considered metaphor conventionality. Unlike Bowdle & Gentner (2005), conventionalization was assessed by sentence-conventionality rating and participants were not forced to answer as fast as possible.

An important methodical question is how can single and double predications account for the difference between categorization and comparison? In the above mentioned studies, double predications were interpreted as indicators of comparison process, since double predications point to properties shared between base and target. However, single predications cannot be directly interpreted as indicators of categorization, since they principally are not a part of base concept's properties and therefore they are not a part of any of base concept's subsets. However, single predications can be considered indicators a

metaphorical concept's distance from the default literal interpretation. When an actual interpretation is different enough from the default literal interpretation, it could behave more "independently" and gain some emergent properties and hence be "more metaphorical". This intuition is somewhat supported by the above mentioned studies. We will therefore interpret single predications as indicators of metaphorical interpretation's semantic distance from default literal interpretation with the assumption that this semantic distance is closely connected (at least in case of metaphors) to the distance (difference) between the number of properties encoded by metaphorical base and target and subsequently to the categorization processing type.

The main hypothesis about novel metaphors in this study was that novel metaphors in categorization form will be interpreted with a significantly higher number of single predications than in comparison form. This is due to novel categorization form metaphors inviting interpretations more distant from default interpretation than novel comparison form metaphors.

About conventional metaphors there were two hypotheses: a) both conventional categorization and comparison sentences will be interpreted similarly, since they both will use the same figurative default concept which serves as a strong attractor. In this case both sentence forms would show similar number of single predications; b) both conventional categorization and comparison form sentences will be interpreted differently due to comparison form inviting interpretations less distant from the original literal interpretation. In this case categorization form sentences would show larger number of single predications.

9.2 Methods

9.2.1 Participants

For this study, 71 participants were recruited (M = 25.6 years, SD = 7.52, 61 females). All participants were undergraduate students of psychology at University of Trnava. In exchange for their participation in the study, students obtained course points.

9.2.2 Stimuli selection

Thirteen raters were originally recruited (M = 26 years, 6 females). All raters were native Slovak speakers and were either master level graduate students or young academics (post-

doc level). Raters were provided with a list of 100 metaphorical sentences. Most of these sentences were selected from current literature and translated to Slovak language, and the rest was constructed by the author of this study. All sentences were in categorization form. Two variables were rated – familiarity and comprehensibility of whole sentences. 9-point Likert scale was used, where 9 represented maximum and 1 minimum of respective variable. Familiarity was used to assess conventionality and in accord with the critique of Thibodeau & Durgin (2011) it was measured on whole sentences and not isolated bases. Comprehensibility ratings were used to control the possible difference in comprehensibility between novel and conventional metaphors which could confound the stimuli and subsequently also the results. Questionnaire containing the sentences was distributed by e-mail individually to every rater in a text document format. There was no time limit and raters were allowed to fill the questionnaire in multiple sessions if they found it too long or too tiring.

Data from three raters were excluded prior to analysis, since they rated no or almost no sentences, including conventional Slovak metaphors, as conventional. The further analysis was based on the data from ten remaining participants (M = 26 years, 4 females). From the original 100 sentences, 16 were selected as stimuli for the experimental phase based on their scores. Half of these sentences was identified as conventional and the other half as novel. Sentences identified as conventional scored 7.4 or higher on familiarity (M = 7.75, SD = 0.32) and 7.9 or higher on comprehensibility (M = 8.28, SD = 0.41). Sentences identified as novel scored 5 or less on familiarity (M = 4.2, SD = 0.82) and 6.8 or higher on comprehensibility (M = 7.35, SD = 0.43). The groups differed significantly in familiarity ($t(11.34) = 9.02$, $p < 0.001$), but also in comprehensibility ($t(4.44) = 13.97$, $p < 0.001$). Comprehensibility may therefore be considered a possible confounding factor in this study. However, it is not entirely unexpected to find novel metaphors as less comprehensible than commonly used conventional metaphors.

All stimuli selected for the experiment are listed in the appendix at the end of this thesis.

9.2.3 Experimental design and procedure

Participants were provided with an online questionnaire with the 16 above mentioned sentences. The task was to paraphrase the base term of every sentence. Participants first read the sentence, for instance *Some ideas are diamonds*, and then they had to complete the second part, beginning for instance with *Some ideas...*, with their own paraphrase.

Participants were not forced to use words *are* or *are like*, because in some sentences it would require production of grammatically rather obscure paraphrases. Participants were allowed to produce paraphrases as long and as elaborated as they wanted. They were also instructed to produce paraphrases with emphasis on the meaning of sentences being paraphrased, rather than on how ‘nice‘ the paraphrases sound or look.

There were two versions of the questionnaire, one with all sentences in categorization form (*X is Y*) and the other with the same set of sentences in comparison form (*X is like Y*). Whether a participant obtained one or the other version of the questionnaire was assessed randomly. 35 participants filled in the categorization form questionnaire and 36 the comparison form questionnaire. In both questionnaires the sentences were presented in the same fixed order and every other sentence was novel. Participants had to fill the questionnaire in one session. There was no time limit and, according to informal reports, it took less than half an hour to complete the questionnaire.

9.2.4 Data preprocessing

Paraphrases were further analyzed by the author. The main variable analyzed was property applicability. When the paraphrased property was applicable on the target only it was considered a single predication and when it was applicable on both target and base it was considered a double predication. When it was not clear whether a property is single or double predication it was rated as ambiguous and excluded from further analysis. Further, one stimulus sentence was excluded prior to predication type analysis. It was the sentence *My house is (like) a museum* which is not, at least in comparison form, possible to be considered a clearly figurative comparison. Both forms of this sentence with all their responses were excluded and further analysis was conducted using responses from the remaining 15 sentences (8 conventional, 7 novel).

After the rating phase, data were statistically analyzed using RStudio.

9.3 Results

Overall, a significantly higher number of double predicates than single predicates was observed ($t(5.76) = 48.161$, $p < 0.001$, $d = 1.66$)). This overall effect was observed also in all combinations of sentence form and conventionality (novel comparison $p = 0.002$; novel categorization $p = 0.02$; conventional comparison $p = 0.006$; conventional categorization $p = 0.03$).

Further, the analysis was focused mainly on single predications. The mean number of single predications was lower for novel comparisons (M = 15.6, SD = 4.5) than for novel categorizations (M = 22.7, SD = 3.35). Conventional comparisons also showed a lower number of single predications (M = 15.6, SD = 9.44) than conventional categorizations (M = 20.75, SD = 8.25). However, as the reader can see, both conventional sentence types and especially conventional comparisons showed large standard deviations (Table 1). Since the datasets of conventional sentences were very small (8 values only), truncation of the datasets was not possible, and winsorization did not show much difference unless it was exceedingly large; therefore data from conventional sentences couldn't be considered reliable. Means with standard deviations are depicted in figure 4.

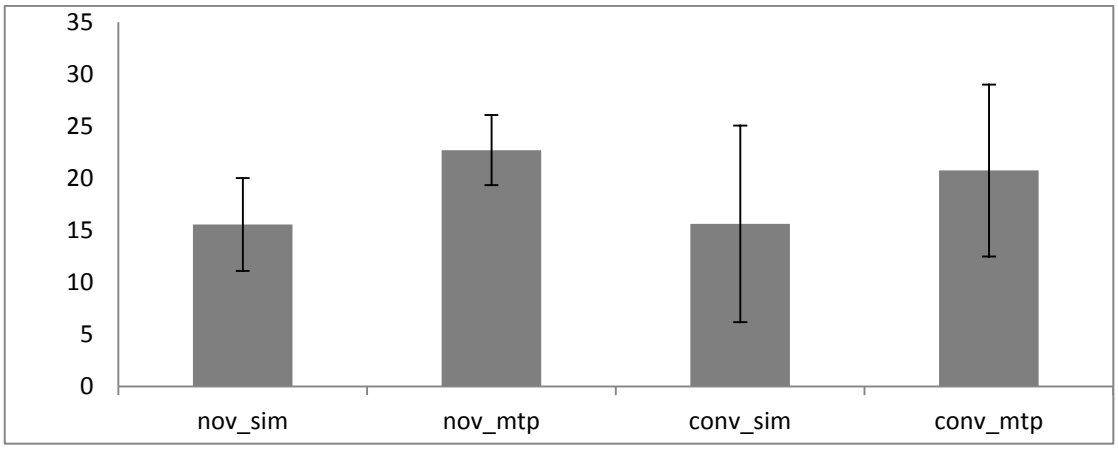

Figure 4 Mean number of single predications for all four conditions with their respective standard deviations (error bars). Abbreviations: nov/conv – novel/conventional, sim/mtp – comparison/categorization form.

Further, interaction of factors conventionality (novel/conventional) and sentence form (comparison/categorization) on sum of single predication properties was analyzed. A 2-way ANOVA did not show any interaction of these factors (F(1, 26) = 0.153, p = 0.7) or effect of the factor conventionality alone (F(1, 26) = 0.138, p = 0.7). This is not surprising given nonhomogeneous data from the conventional sentences. The only significant factor was sentence form (F(1, 26) = 5.57, p = 0.02). T-tests also confirmed this; when applied selectively, the only significant difference was observed between the sums of single predications as function of sentence form in novel metaphors (t(3.38) = 11.131, p = 0.006, d = 2.03), but not in conventional metaphors (t(1.15) = 13.757, p = 0.2). Notice that in novel metaphors the effect size was very large. These data are therefore in accord with the

hypothesis that novel categorization sentences will show significantly more single predication properties than novel comparisons.

There was no direct effect of conventionalization on single predication properties in categorization ($t(0.61) = 9.492$, $p = 0.5$) or comparison form ($t(-0.014) = 10.257$, $p = 0.9$). No differences were observed in the number of double predications in any condition.

Since the conventional metaphors showed a large variance, the results from every sentence were normalized to the average of each sentence regardless of the sentence form. We have used the one-sample t-tests to examine the differences between sentences presented in comparison and categorization form. Similarly to the findings reported above, results showed a significant difference in the number of single predications between the two forms in novel metaphors ($t(2.97) = 6$, $p = 0.02$) but not in conventional metaphors ($t(1.97) = 7$, $p = 0.08$). The normalization did not solve the problem with large standard deviations observed in conventional metaphors, therefore our conclusions (or their absence) about these sentences remain unchanged.

9.4 Interpretation

The results of this study confirm one of the hypotheses. We showed that novel metaphors in categorization form are more often interpreted with single predications than novel metaphors in comparison form. This supports the view that interpretations of novel comparisons are less distant from the default interpretation than novel categorizations, and it is in accord with the theoretical account presented in the previous chapter.

Unfortunately, the data from conventional sentences could not be considered reliable, therefore the hypotheses proposed about conventional metaphors remain open and it is not possible, based on our results, to determine whether interpretations of conventional categorizations and comparisons are similar or not.

9.5 Limiting factors

There are few possible limiting factors. First, a small number of values was present in analyzed datasets. Unfortunately, this is a common constraint in this type of studies, since it is not possible to provide participants with a very large number of sentences (for practical reasons such as difficulty and duration of these questionnaires). Another aspect that may be considered problematic is stimulus selection – in the rating phase only the

categorization form sentences were used, and therefore ratings did not cover comparison form sentences. Finally, the variance observed in conventional metaphors interpretation remains unexplained, but it probably is a result of variation in natural language.

Conclusion

The main purpose of this work was to evaluate current theories of metaphor and to try to solve the long lasting problem of whether and when are metaphors processed as comparisons or categorizations. We focused mainly on the career of metaphor theory, the class-inclusion theory and also on the relevance theory. We tried to integrate these, sometimes rather contradictory, views on metaphor in order to solve the main problem.

Based on the above mentioned theories and the study of Tenenbaum & Xu (2000) we proposed that it is the difference between the total number of properties of metaphorical base concept and literal target concept what determines whether a metaphor will be perceived as comparison or categorization - the larger the difference, the higher the chance that the metaphor will be processed as a comparison.

We tested this proposal experimentally using the free paraphrase method. We provided participants with novel and conventional metaphorical utterances in categorization or comparison form and observed the properties they used to paraphrase these metaphors. Two property types were recognized – double predications which can be used to describe both base and target of a sentence and single predications which can be used to describe target only. We considered single predications to be indicative of greater difference between the concepts than double predications. Based on this, we hypothesized that single predications will be more frequently used to describe novel categorization form metaphors than novel comparison form metaphors.

About conventional metaphors, we proposed two incompatible hypotheses. We expected conventional sentences in both forms to be interpreted with an equal number of single predications due to conventional metaphorical interpretation being a strong attractor. Alternatively we expected that, similarly to novel metaphors, the sentence form will determine the interpretation type and therefore again, categorizations will be interpreted using single predications more often than comparisons. Unfortunately, our data from conventional metaphors could not be considered reliable, therefore we were able to draw conclusions only about the novel sentences. In accord with the hypothesis, novel categorizations were more often interpreted using single predications than novel comparisons.

Additionally, we were able to use our theoretical integration to provide a simple description of aberrant semantic processing in autism and schizophrenia (of course with emphasis on metaphor comprehension). We based our view on neuroscientific findings, mainly from the subfields of microanatomy and cognitive neuroscience. Based on this evidence we described autism in terms of overfitting and schizophrenia in terms of underfitting. This allowed us to explain autism as being too restrictive and thus not allowing for much variability of words' interpretations (which is probably required for novel metaphor comprehension) and schizophrenia as being too nonrestrictive and therefore facilitating more diffuse semantic processing.

Whether the propositions formulated in this thesis are valid or not, will be determined by future theoretical as well as experimental work.

References

Aalst, Rubin, Verbeek, Dongen, Kindler & Günther (2010). Process mining: a two-step approach to balance between underfitting and overfitting. *Software & Systems Modeling* 9(1), pp. 87-111.

Aisenman (1999). Structure-Mapping and the Simile-Metaphor Preference. *Metaphor and Symbol* 14(1), pp. 45-51.

Aristotle (2004). *Rhetoric* (transl. W. Rhys Roberts). Dover Publications.

Bakouie, Zendehrouh, Gharibzadeh (2009). Does a Kind of Over-Fitting Occur in the Brain of Autistic Patients? *Journal of Neuropsychiatry and Clinical Neuroscience* 21(3), p. 343.

Barnden (2012). Metaphor and Simile: Fallacies Concerning Comparison, Ellipsis, and Inter-Paraphrase. *Metaphor and Symbol* 27(4), pp. 265-282.

Bowdle & Gentner (1997). Informativity and Asymmetry in Comparisons. *Cognitive Psychology* 34(3), pp. 244-286.

Bowdle & Gentner (1999). Metaphor Comprehension: From Comparison to Categorization. *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society*, pp. 90-95.

Bowdle & Gentner (2005). The Career of Metaphor. *Psychological Review* 112(1), pp. 193-216.

Buxhoeveden, Semendeferi, Buckwalter, Schenker, Switzer & Courchese (2006). Reduced minicolumns in the frontal cortex of patients with autism. *Neuropathology and Applied Neurobiology* 32(5), pp. 483-491.

Buxhoeven, Switala, Litaker, Roy & Casanova (2001). Lateralization of minicolumns in human planum temporale is absent in nonhuman primate cortex. *Brain, Behavior and Evolution* 57(6), pp. 349-358.

Casanova (2007). The Neuropathology of Autism. *Brain Pathology* 17(4), pp. 422-433.

Catani, Jones & Ffytche (2005). Perisylvian Language Network of the Human Brain. *Annals of Neurology* 57(1), pp. 8-16.

Cruys, de-Wit, Evers, Boets & Wagemans (2013). Weak priors versus overfitting of predictions in autism: Reply to Pellicano and Burr (*TICS*, 2012). *i-Perception* 4(2), pp. 95-97.

Dolník (2013). Všeobecná jazykoveda [*General linguistics*] (2nd edition). Bratislava: VEDA, vydavateľstvo Slovenskej akadémie vied.

Faust & Kenett (2014). Rigidity, chaos and integration: hemispheric interaction and individual differences in metaphor comprehension. *Frontiers in Human Neuroscience* 8.

Faust & Mashal (2007). The role of the right cerebral hemisphere in processing novel metaphoric expressions taken from poetry: A divided visual field study. *Neuropsychologia* 45(4), pp. 860-870.

Fletcher, Whitaker, Tao, DuBray, Froehlich, Ravichandran, Alexander, Bigler, Lange & Lainhart (2010). Microstructural connectivity of the arcuate fasciculus in adolescents with high-functioning autism. *NeuroImage* 51(3), pp. 1117-1125

Gannon, Holloway, Broadfield & Braun (1998). Asymmetry of chimpanzee planum temporale: humanlike pattern of Wernicke's brain language area homolog. *Science* 279(5348), pp. 220-222.

Gentner & Grudin (1985). The Evolution of Mental Metaphors in Psychology: A 90-Year Retrospective. *American Psychologist* 40(2), pp. 181-192.

Gentner & Markman (1997). Structure Mapping in Analogy and Similarity. *American Psychologist* 52(1), pp.45-56.

Gentner (1982). Are Scientific Analogies Metaphors? In Miall (Ed.) *Metaphor: Problems and perspectives* (pp. 106-132). Harvester Press.

Gentner (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science* 7(2), pp. 155-170.

Gentner, Bowdle, Wolff & Boronat (2001). Metaphor is like analogy. In Gentner, Holyoak & Kokinov (Eds.), *The analogyical mind: Perspectives from cognitive science* (pp. 199-253). Cambrige MA, MIT Press.

Gibbs, Lima & Francozo (2004). Metaphor is grounded in embodied experience. *Journal of Pragmatics* 36(7), pp. 1189-1210.

Glucksberg & Haught (2006). On the Relation Between Metaphor and Simile: When Comparison Fails. *Mind & Language* 21(3), pp. 360-378.

Glucksberg (2001). *Understanding Figurative Language: From Metaphors to Idioms.* Oxford University Press.

Glucksberg (2003). The psycholinguistics of metaphor. *TRENDS in Cognitive Science* 7(2), pp. 92-96.

Gold, Faust & Goldstein (2010). Semantic integration during metaphor comprehension in Asperger syndrome. *Brain & Language* 113(3), pp. 124-134.

Hasan, Kremer, Gruber, Schneider-Axmann, Guse, Reith, Falkai & Wobrock (2011). Planum temporale asymmetry to the right hemisphere in first-episode schizophrenia. *Psychiatry Research: Neuroimaging* 193(1), pp. 56-59.

Hickok (2009). The functional neuroanatomy of language. *Physics of Life Reviews* 6(3), pp. 121-143.

Hoppenbrouwers, Vandermosten & Boets (2014). Autism as a disconnection syndrome: A qualitative and quantitative review of diffusion tensor imaging studies. *Research in Autism Spectrum Disorders* 8(4), pp. 387-412

Hossain (2013). Between concepts and metaphor: reviewng Nietzeches doctrine of truth. *International Journal of Philosophy* 1(1), pp. 6-20.

Hutsler & Galuske (2003). Hemispheric asymmetries in cerebral cortical networks. *TRENDS in Neurosciences* 26, pp. 429-435.

Chance (2014). The cortical microstructural basis of lateralized cognition: a review. *Frontiers in Psychology* 5, pp. 1-8.

Chance, Casanova, Switala & Crow (2006). Minicolumnar structure in Heschl's gyrus and planum temporale: asymmetries in relation to sex and callosal fiber number. *Neuroscience* 143(4), pp. 1041-1050.

Chance, Casanova, Switala & Crow (2008). Auditory cortex asymmetry, altered minicolumn spacing and absence of ageing effects in schizophrenia. *Brain* 131(12), pp. 3178-3192.

Chance, Sawyer, Clover, Wicinski, Hof & Crow (2013). Hemispheric asymmetry in the fusiform gyrus distinguishes *Homo sapiens* from chimpanzees. *Brain Structure and Function* 218(6), pp. 1391-1405.

Chance, Walker & Crow (2005). Reduced density of calbindin-immunoreactive interneurons in the planum temporale in schizophrenia. *Brain Research 1046(1-2),* pp. 32-37.

Chiappe & Kennedy (2000). Are Metaphors Elliptical Similes? *Journal of Psycholinguistic Research* 29(4), pp. 371-398.

Johnston, Rossell & Gleeson (2008). Evidence of semantic processing abnormalities in schizotypy using an indirect semantic priming task. *Journal of Nervous and Mental Disease* 196(9), pp. 694-701.

Jones & Estes (2006). Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language* 55(1), pp. 18-32.

Jung-Beeman (2005). Bilateral brain processes for comprehending natural language. *TRENDS in Cognitive Sciences* 9, pp. 512-518.

Kenett, Anaki & Faust (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience* 8.

Keysar, Shen, Glucksber & Horton (2000). Conventional Language: How Metaphorical Is It? *Journal of Memory and Language* 43, pp. 576-593.

Kiefer, Weisbrod, Kern, Maier & Spitzer (1998). Right Hemisphere Activation during Indirect Semantic Priming: Evidence from Event-Related Potentials. *Brain and Language* 64(3), pp. 377-408.

Kövecses (2010). *Metaphor: A Practical Introduction* (2nd Edition). Oxford University Press.

Lakoff (1993). The contemporary theory of metaphor. In Ortony (Ed.) *Metaphor and Thought* (2nd edition). Cambridge University Press.

Leroux, Delcroix & Dollfus (2014). Left fronto-temporal dysconnectivity within the language network in schizophrenia: An fMRI and DTI study. *Psychiatry Research: Neuroimaging* 223(3), pp. 261-267.

Oertel, Knöchel, Rotarska-Jagiela, Schönmyer, Linder, Ven, Haenschel, Uhlhaas, Maurer & Linden (2010). Reduced Laterality as a Trait Marker of Schizophrenia – Evidence from Structural and Functional Neuroimaging. *The Journal of Neuroscience* 30(6), pp. 2289-2299.

Ortony (1979). Beyond Literal Similarity. *Psychological Review* 86(3), pp. 161-180.

Ortony (1993). Metaphor, language, and thought. In Ortony (Ed.) *Metaphor and Thought* (2nd edition). Cambridge University Press.

Pellicano & Burr (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in Cognitive Sciences* 16(10), pp. 504-510.

Pierce & Chiappe (2009). The Roles of Aptness, Conventionality, and Working Memory in the Production of Metaphors and Similes. *Metaphor and Symbol*, 24(1), pp. 1-19.

Sætrevik & Specht (2012). Cognitive conflict in asyllable identification task causes transient activation of speech perception area. *Brain and Cognition* 78(3), pp. 200-205.

Selemon & Goldman-Rakic (1999). The Reduced Neuropil Hypothesis: A Circuit Based Model of Schizophrenia. *Biological Psychiatry* 45(1), pp. 17-25.

Shukla, Keehn, Smylie & Müller (2011). Microstructural abnormalities of short-distance white matter tracts in autism spectrum disorder. *Neuropsychologia* 49(5), pp. 1378-1382.

Scheibel (1988). Dendritic correlates of human cortical function. *Archives Italiennes de Biologie* 126(4), pp. 347-375.

Scheibel, Paul, Fried, Forsythe, Tomiyasu, Wechsler, Kao & Slotnick (1985). Dendritic organization of the anterior speech area. *Experimental Neurology* 87(1), pp. 109-117.

Schenker, Sherwood, Hof & Semendeferi (2007). Microstructural asymmetries of the cerebral cortex in humans and other animals. *Special Topics in Primatology* 5, pp. 82-118.

Spitzer, Braun, Maier, Hermle & Maher (1993). Indirect semantic priming in schizophrenic patients. *Schizophrenia Research* 11(1), pp. 71-80.

Srivastava, Hinton, Krizhevsky, Sutskever & Salakhutdinov (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research* 15, pp. 1929-1958.

Steen, Mull, McClure, Hamer & Lieberman (2006). Brain volume in first-episode schizophrenia. *British Journal of Psychiatry* 188, pp. 510-518.

Stephane, Pellizzer, Fletcher & McClannahan (2007). Empirical evaluation of language disorder in schizophrenia. *Journal of Psychiatry and Neuroscience* 32(4), pp. 250-258.

Stern (2008). The Life and Death of a Metaphor, or the Metaphysics of Metaphor. *The Baltic International Yearbook of Cognition, Logic and Communication* 3, pp. 1-43.

Tendahl & Gibbs (2008). Complementary perspectives on metaphor: Cognitive linguistics and relevance theory. *Journal of Pragmatics* 40, pp. 1823-1864

Tenenbaum & Xu (2000). Word learning as Bayesian inference. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, pp. 517-522.

Thibodeau & Durgin (2011). Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol* 26(3), pp. 206-226.

Toga & Thompson (2003). Mapping brain asymmetry. *Nature Reviews Neuroscience* 4(1), pp. 37-48.

Tversky (1977). Features of similarity. *Psychological Review* 84(4), pp. 327-352.

Upadhyay, Hallock, Ducros, Kim & Ronen (2007). Diffusion tensor spectroscopy and imaging of the arcuate fasciculus. *NeuroImage* 39(1), pp. 1-9.

Utsumi & Sakamoto (2010). Predicative metaphor comprehension as indirect categorization. *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, pp. 1034-1039.

Vulchanova, Saldaña, Chahboun & Vulchanov (2015). Figurative language processing in atypical populations: the ASD perspective. *Frontiers in Human Neuroscience* 9.

Watkins, Paus, Lerch, Zijdenbos, Collins, Neelin, Taylor, Worsley & Evans (2001). Structural Asymmetries in the Human Brain: a Voxel-based Statistical Analysis of 142 MRI Scans. *Cerebral cortex* 11(9), pp. 868-877.

Wilson & Sperber (2006). Relevance Theory. In: Horn & Ward (eds.), *Handbook of Pragmatics* (pp. 249-287). Wiley-Blackwell.

Wilson (2011). Parellels and differences in the treatment of metaphor in relevance theory and cogntive linguistics. *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 128, pp. 195-213.

Wolff & Gentner (1992). The Time Course of Metaphor Comprehension. *Proceedings of the fourteenth annual conference of the Cognitive Science Society*, pp. 504-509.

Zeev-Wolf, Goldstein, Levkovitz & Faust (2014). Fine-coarse processing in schizophrenia: A reversed pattern of hemispheric dominance. *Neuropsychologia* 56, pp. 119-128.

Zharikov & Gentner (2002). Why do metaphors seem deeper than similes? *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, pp. 976-981.

Appendix

Stimuli used in our study with their respective ratings of comprehensibility and familiarity. Conventional and novel stimuli are listed in separate tables. Note that English translations are only approximate and don't necessarily express the same connotations as the original Slovak versions.

A. Conventional metaphors

	comprehensibility	**familiarity**
Tá modelka je špáratko. (That fashion model is a stick)	8.9	7.4
Ten chirurg je mäsiar. (That surgeon is a butcher)	8.7	8.1
Sklamanie je rana. (Dissapointment is a wound)	8.6	7.4
To dieťa je anjel. (That baby is an angel)	8.4	8.2
Myseľ je počítač. (Mind is a computer)	8	7.5
Čakanie v rade je smrť. (Waiting in line is death)	7.9	7.6
Život je cesta. (Life is a journey)	7.9	7.9
Planéta Zem je matka. (Planet Earth is mother)	7.9	7.9

B. Novel metpahors

	comprehensibility	**familiarity**
Niektoré myšlienky sú diamanty. (Some ideas are diamonds)	8.0	5.0
Rešpekt je klenot. (Respect is a gem)	7.8	4.5
Dav je veľká rieka. (Crowd is a big river)	7.5	5.0
Úsmev je veľvyslanec. (Smile is an ambassador)	7.5	3.6
Môj dom je múzeum. (My house is a museum)	7.3	4.7
Zlepšovanie sa je kvitnutie. (Improvement is a flowerage)	7.0	3.6
Šéfove slová sú tesáky. (Boss's words are sharp teeth)	6.9	2.7
Motýle sú tanečníci. (Butterflies are dancers)	6.8	4.6