Scalar Implicatures and Existential Import: Experimental Study on Quantifiers in Natural Language

MSc Thesis (*Afstudeerscriptie*)

written by

Maria Spychalska
(born November 21, 1980 in Toruń, Poland)

under the supervision of **Prof. Dr Michiel van Lambalgen**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
August 31, 2009	Dr Maria Aloni
	Dr Robert van Rooij
	Prof. Dr Jouko Väänänen
	Prof. Dr Frank Veltman

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

To Merrien Ar-Imrahillis,

with a promise of re-constructing her world.

Contents

Acknowledgments

No individual is a deserted, lonely monad. Sharing knowledge, discussing ideas, helping each other – those are foundations of a good academic work. I would like to thank to many people, without whom this thesis would have never been written or would not have had the present shape.

To my family – my parents and my sister Anna, for their support, motivation, love and help at every single moment would I need it, I want to thank at the first place. Special thanks I owe to my mother who typed in all the 177 tests from the second experiment, which makes 10620(!) 0's and 1's. If not of her help I would have not managed to finish my thesis on time, simply because having my hand broken, I was limited in using computer. My sister had a special share in the thesis-creation in writing for me a "daily working plan" at the moment of despair after I had lost the source files of the first version of the thesis. Last but not least, I would like to thank to my father for comments on experimental materials and interesting remarks about the language.

I am particularly grateful to Marcin Mostowski, who was the first to suggest me investigating syllogistic with "most" together with the problem of existential import and empty domains. Here, I need to mention also Paula Quinon (neé Bucholc), whom I want to thank for our talks about syllogistic and empty terms. I would like to emphasize that her master thesis was the first inspiration for my work.

Two people contributed directly to my thesis. I mean Piotr Wilkin, who is the co-author of the idea that existence is a modal index. The present shape of Chapter 4 is to big extent a result of our long discussions about possible worlds, elves and semantics. The second person is Eryk Kopczyński, who wrote a computer program for generating syllogisms with "most" and made for me also a small "test-mixer", so a program randomizing the order of questions in tests. Since programming has always been (and probably it will stay so) my weakness, if not of Eryk's help I would probably end up generating syllogisms by hand. Finally, both Eryk and Piotr read my formalization of vague quantifiers helping me to polish my ideas and the style of writing about logic. I would like to express my appreciation for this.

For useful comments on my text – its style and content, remarks on experimental materials, discussing ideas, and more technical help (with latex, graphs, tables, etc.) I want to thank to: Eryk, Jakub Szymanik (especially for talks about experiments and hints about style of writing and edition), Jarmo Kontinen (especially for deleting all my excellent but irrelevant metaphors, adjectives, comparisons and other literary inclusions), Justyna Grudzińska, Marek Czarnecki, Nina Gierasiumczuk (special thanks for the being my "picture-emergency"), Piotr, Rafał Olszewski and Tadeusz Ciecierski. To Jarmo I am particularly grateful also for support, when I lost my data. I appreciate this a lot, because not being left totally alone in this misfortune, I was able to find motivation for further work. Finally, I thank to Jonathan Zvesper, who agreed to proofread my thesis, for excellent language corrections and other interesting comments on text.

I cannot forget about people who helped me make my first steps in logic. I mean my teachers: Teresa Hołówka (she was the first to make me aware of the fact that logic in common reasoning may differ from what logicians in their offices would expect), Marcin Mostowski and Cezary Cieślinski. Special thanks also to Konrad Zdanowski, who devoted many hours in Warsaw for making logic (especially model theory) less mysterious for me; and to my friends Eryk and Piotr, who were always willing to reply my most weird logic questions, usually with a socratic method, hoping that I can find the answer myself.

The time I spend in Amsterdam at the ILLC was probably the most scientifically inspiring for me in my whole life. Especially important was for me meeting Michiel van Lambalgen, who became my thesis supervisor. His attitude to logic, philosophy and reasoning strongly influenced my thinking and I am sure that I will later continue research on human reasoning in this spirit. I am grateful to Michiel for long and very interesting philosophical discussions, reading my thesis, for important hints and remarks, and for showing me the path leading from logic, through philosophy, to psychology and back.

I would like to thank also to all the committee members: Frank Veltman, Jouko Väänänen, Maria Aloni and Robert van Rooij for agreeing to read my work and participate in the defence. Special thanks I owe to Jouko Väänänen for very useful remarks about generalized quantifiers and to Robert van Rooij for being my mentor for two years and helping me with my my first study-plan. I take also the opportunity to mention Yurii Khomskii and thank him for afternoons he spent explaining us, students, mysteries of recursion theory. I would like to thank to Henk Zeevat and Reinhard Blutner for comments on my research in Szklarska Poręba Workshop, and to Benedikt Loewe for help in solving all the small study-problems. Last but not least I thank to Tanja Kassenaar and Peter van Ormondt – for patiently replying all my emails with strange questions about things I probably should know myself.

I need to emphasize that participating in the *Master of Logic Programme* would have not been possible for me, if I had not gotten the financial support. I am especially grateful to the *Beth Foundation* and to the people that had decided for awarding me the scholarship. I would not have come to Amsterdam if I had not gotten this scholarship. I thank also to *Nuffic* organization for awarding me on the later stage the *Huygnes* scholarship, which opened for a me a possibility of staying in Amsterdam for two instead of a one year.

The experimental study was a kind of adventure for me, especially that I had first to learn (from the very basics) how to do and report statistical analyzes. I want to thank to Conor Dolan for excellent teaching and useful hints, to Marek Młodożeniec for evaluating my first independent analyses, to Marcin Zajenkowski for additional comments and to Andy Field for writing the best statistics handbook one may imagine, which was my guide throughout whole my work.

Many people were my first "experimental victims" and their comments helped me considerably polish the experimental materials or became the basis of qualitative evaluation of results. I mean here my parents, my sister, my friends: Danuta Kozakiewicz, Irena Grudzińska, Jakub Abramczuk, Marta Urbańska, and other friends of friends, whose names I even do not know or do not remember. I am very grateful to people that conducted the experiments or agreed to "lend me" part of their lectures together with their students for this purpose: Agnieszka and Dariusz Dąbrowscy, Bogdan Dziobkowski, Joanna Golińska-Pilarek, Justyna, Michał Sala, Tadeusz and Tomasz Wiśniewski. I thank to friends from the *Orkon* group for replying and commenting on our forum my strange questions about existence of mermogliny and their pointed ears; last but not least to all anonymous students who agreed to be tested in the experiments. Their role in this research cannot be overstated.

Finally, I want to thank to people who helped me find my place here: Andreas Witzel, Amélie Gheerbrant, Danka (especially for "pushing" me and encouraging to apply for the Programm), Gaëlle Fontaine (she was the first person to drink with me coffee in Amsterdam on the day I arrived and felt completely lost, although she did not know me at all at that moment – thanks!), Jakub & Nina (for being my little Warsaw here), Jarmo (kiitos!), Lauri Keskinen, Michał Łukasiewicz, Pablo Cubides Kovacsics (for my first and hopefully last travel on bike's handle bars!), Rogier Jacobsz (for common struggle to understand model theory), Sophia Katarenko, Stephan Schroevers (for explaining me the complicated aspects of Dutch culture), Szymon Klarman, Theodora Achourioti, and other colleagues from the Institute. Thank to them my homesickness was easier to bear and finally changed even into summer "Amsterdam-sickness". I am sure that I will always remember the time I spent in Amsterdam with sentiment and that I will be coming back to this romantic city with pleasure, since...

Dans le port d'Amsterdam
Y a des marins qui chantent.

Introduction

This thesis explores two related themes: *scalar implicatures* of the quantifiers "most" and "some" and the philosophical problem of *existential import*. Traditionally, existential import refers to the question whether the universal categorial sentence "All A's are B" implies that there are any A's. The origin of this question lies in the "subalternation" relation of Aristotle's Square of Opposition, i.e. in the inference:

$$\frac{\text{All A's are B}}{\text{Some A's are B}}$$

Some have argued that since "some" is existentially loaded (it is represented in predicate logic with the use of the existential quantifier), then also "all" has to be, otherwise we would have to lose the above inference.

Scalar implicature is a pragmatic property, based on one of the Gricean conversational maxims, i.e. on the Maxim of Quantity, which is a commandment to be as informative as possible. The use of the quantifier "most" or "some" in a sentence implicates that all similar utterances using an informationally stronger term are not true. Since "all" is supposed to be stronger than "most" and "most" stronger than "some", both "some" and "most" implicate the negation of "all", and "some" implicates also the negation of "most".

The first step of our study was to extend the language of Aristotelian syllogistic by adding the quantifier "most", and analyze all the possible inferences that can be generated in this way. The purpose of this was to investigate the cognitive aspect of syllogisms with "most" – their (cognitive) difficulty with comparison to the traditional ones, existential import of this quantifier and its implicature.

The idea of formalizing natural reasoning with the quantifier "most" is not completely new. The first syllogism with "most" was given by Rescher (Rescher, 1962) and is known as "Rescher's syllogism":

$$\frac{\begin{array}{c}\text{Most A-s are B}\\ \text{Most A-s are C}\end{array}}{\text{Some B-s are C}}$$

An interesting approach was also presented by Chater and Oaksford (1999), where authors developed a probability heuristic model for syllogistic reasoning, extending it also to generalized quantifiers such as "most" and "few".[1]

Syllogistic with "most" was hence a starting point for our study. Since no comprehensive research in this field has been done so far, we found this direction worth exploring and the gap worth filling. Analysis of complex, two-premise reasoning required, however, preliminary examining how "most" works in simple one-premise inferences. This stage turned out to be very interesting and allowing many observations. Finally the more detailed study on syllogistic had to be left for further research and, although we devoted some time to analysis of two-premise inferences, our work focused more on simple direct reasoning with "most". We were interested in classically valid inferences on one hand, e.g. inferences of "Some A's are B" from "Most A's are B" compared with similar inferences with "All A's are B" being a premise – the latter requiring for validity the additional assumption that domains are non-empty. Thus we compared the existential import of "most" with that of "all". On the other hand we investigated pragmatic inferences, i.e. scalar implicatures, such as "Most A's are B" implicating "Not all A's are B", and "Some A's are B" implicating "Not all A's are B" (resp. "Not Most A's are B").

We conducted a series of experiments to see what kind of inferences people are willing to make with those quantifiers. The first three experiments focused on checking so-called *active logical competence*, so people's readiness to *generate* inferences for given premises. We checked the semantics of "most" in small domains with the use of simple picture-tasks, direct inferences with premises with "most", as well as with Aristotelian quantifiers, and finally syllogisms with 'most". The text-tests were conducted in two groups, with tasks using so-called empty (non-referring) and non-empty terms. The purpose of this was to check whether the ontological status of the entities referred to has any influence on the inferences which require additional assumption of non-emptiness of domains. The second series of experiments was designed to check *passive logical competence*, so people's evaluation of inferences as correct. In this part we considered only direct inferences. Whereas the first experiment has to be treated as a more qualitative study – the responses given by subjects required preliminary interpretation and also the character of the experiment prevented from asking many questions in one test, the second one was carefully planned for detailed statistical analyses and quantitative testing of hypotheses. For this reason it allows more ceratin conclusions and is more interesting with respect to significance of results.

Based on our results we propose modeling scalar implicatures of quantifiers "some" and "most" in terms of fuzzy semantics. We describe these quantifiers

[1]It is doubtful whether "few", as an expression with not determined cardinality semantics, is indeed a generalized quantifier and not just an adjective – with semantics dependent on linguistic context and not on the cardinality of the domain. Later we will see that "some" meets similar problems.

as vague with flexible default denotations. Roughly speaking it means that those quantifiers have some default meanings, which can be changed (i.e. extended) in the process of reasoning. We propose that the implicature "not all" is a part of the default meaning of both "some" and "most". However, since the denotations are flexible, both considered quantifiers are extendable to the case "all".

Chapter 4 is devoted wholly to the philosophical problem of existential import. Inspired by our empirical results showing that people's treatment of fictional sentences containing quantifiers is invariant on whether the sentence mentions fictional or real objects, we propose a unified approach to existential presuppositions that is valid for both imaginary and actual content. We use possible worlds semantics, treating worlds as mental constructions describing certain world states. Within that framework, we analyze sentences about the real world as referring to the real world description, treated as a special case of "possible world", and existential sentences as applying a modal index restricting the domain of discourse to a certain world or class of worlds. We analyze the empirical data concerning quantifier usage with this theory, giving an account of how existential presuppositions work in fictional discourse.

In the last chapter we briefly summarize our main results and pose some open questions to answer in the prospective research.

Chapter 1

Preliminaries

1.1 Aristotelian syllogistic

Aristotle's syllogistic is the oldest formalized logical calculus and was invented as a device of correct reasoning. Syllogistic was a term logic – it was neither propositional, nor predicate logic – and focused precisely on the analysis of quantification, although the notion of a quantifier appeared much later.

The language of Aristotle's syllogistic contains two groups of symbols:

1. *Termini* (*terms*): $\mathbf{T} = \{A, B, C\}$ [1]
2. *Copulae* (*connectives*): $\mathbf{Q} = \{$All, Some, No, Some not$\}$

Contemporarily, *copulae* are called "Aristotelian quantifiers" and treated as $\langle 1, 1\rangle$ generalized quantifiers: all, some, no, not all (some not).

Terms and connectives create together *categorial sentences*. Each categorial sentence contains a subject (*Subjectum*) and a predicate (*Praedicatum*) bound by a copula. Thus a categorial sentence is a sentence of the following form: $Q_i(X, Y)$, where $Q_i \in \mathbf{Q}$, and $X, Y \in Termini$ and $X \neq Y$, X is a subject and Y is a predicate.[2] There are four basic categorial sentences:

1. A - $All(A, B)$ meaning *All A's are B* (*universal affirmative*).
2. I - $Some(A, B)$ meaning *Some A's are B* (*particular affirmative*).
3. E - $No(A, B)$ meaning *No A's are B* (*universal negative*).

[1] Traditionally: S, M, P. These letters were used for mnemotechnic reasons. Throughout we will use A, B, C.

[2] Traditionally a predicate is mentioned first and a subject subsequently, which was connected with the specificity of the Greek language. We use the opposite – more natural in English – way of reading.

4. O - *Somenot*(A, B) meaning *Some A's are not B* (*particular negative*).[3]

The letters A, I, E, O refer to the traditional names of the categorial sentences.

Categorial sentences create the *Square of Opposition*, which is a group of these embodied in a square diagram. (See Fig. 1.1.) The following dependencies make up the *Square*:

- *All*(A, B) and *Somenot*(A, B) are contradictories.
- *No*(A, B) and *Some*(A, B) are contradictories.
- *All*(A, B) and *No*(A, B) are contraries.
- *Some*(A, B) and *Somenot*(A, B) are subcontraries.
- *Some*(A, B) is a subaltern of *All*(A, B).
- *Somenot*(A, B) is a subaltern of *No*(A, B).

We need still to explain the notions. Thus let ϕ and ψ be sentences:

Definition 1. *ϕ and ψ are contradictories if and only if they cannot both be true and they cannot both be false.*

Definition 2. *ϕ and ψ are contraries if and only if they can both be false, but they cannot both be true.*

Definition 3. *ϕ and ψ are subcontraries if and only if they can both be true, but they cannot both be false.*

Definition 4. *ϕ is a subaltern of ψ if and only if it cannot be the case that ψ is true and the subaltern ϕ is false.*

1.1.1 Syllogisms

Syllogisms are tuples of categorial sentences, where one of the sentences is a conclusion and the rest are premises. We differentiate between one-premise direct inferences and two-premise syllogisms. Actually only two-premise arguments are traditionally called syllogisms. Direct inferences are nonetheless important since proper i.e. two-premise syllogisms base on them.

Definition 5. *A direct inference is an ordered pair of categorial sentences:*
$\langle \phi = Q_i(X, Y), \psi = Q_j(X, Y) \rangle$, *where ϕ is a premise and ψ is a conclusion.*

[3] Throughout we use *Somenot* to denote the quantifier "some not".

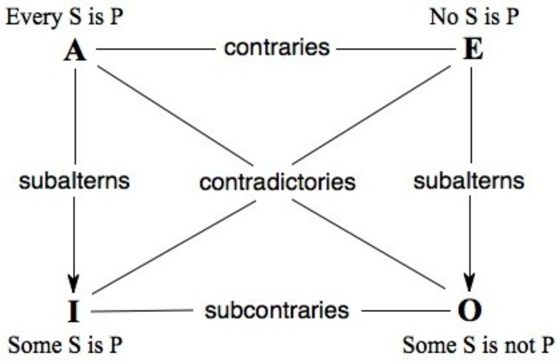


Figure 1.1: The Traditional Square of Opposition

Definition 6. A syllogism *is a triple of categorial sentences of the form:*
$\langle \phi_1 = Q_i([Y,X]^*), \phi_2 = Q_j([Z,Y]^*), \psi = Q_k(X,Z)\rangle$, *where:* ϕ_1, ϕ_2 *are premises and* ψ *is a conclusion,* Z *is a* major term *(predicate in the conclusion),* X *is a* minor term *(subject in the conclusion),* Y *is a* middle term, *and* $[X,Y]^*$ *means that terms can be permuted within the sentence.*

Accordingly a premise with a major term is called a major premise *and a premise with a minor term is called a* minor premise.

It is quite obvious why not all possible triples of all possible categorial sentences create "syntactically" correct syllogisms. For example, we would not like a syllogism of the following form: $\langle Q_i(A,B), Q_j(B,A), Q_k(B,C)\rangle$, since there is no connection between the premises – there is no "middle term" and the term C does not occur in the premises at all.

There are four possible patterns for syntactically correct syllogisms. They are so-called *figures* and depend on the arrangement of terms.[4]

$$\begin{array}{cccc} Fig.I & Fig.II & Fig.III & Fig.IV \\ \\ BA & AB & BA & AB \\ CB & CB & BC & BC \\ \hline CA & CA & CA & CA \end{array}$$

Filling the figures with quantifiers from the set $\mathbf{Q}$, we obtain *syllogistic moods.* There are $4^3 * 4 = 256$ of all possible moods, but only some of them are logically valid inference patterns – namely semantically valid (true in all possible models).

Definition 7. *Let* $S = \langle \phi_1, \phi_2, \psi\rangle$ *be a syllogism.* S *is semantically valid if and only if whenever* ϕ_1 *and* ϕ_2 *are true,* ψ *is also true.*

[4]The order of terms in the given figures is the same as in Aristotle's, however traditionally it should be read backwards. We use further the traditional figures reading them from right to left, which makes no difference since terms in syllogisms can be permuted.

1.1.2 The problem of empty domains

Traditionally it is claimed that there are 24 valid Aristotelian syllogisms. According to the modern view, this is true only with the additional assumption that terms used in the syllogisms are non-empty. Contemporarily, merely 5 of Aristotelian syllogisms are accepted as inferences correct in an "absolute sense", i.e. without any assumption about the domains. These are moods with no particular conclusion that is inferred from universal premises.[5] For example from the following premises:

$$\mathit{All}(B, A)$$
$$\mathit{All}(C, B)$$

one can infer two conclusions:

$$\mathit{All}(C, A) \tag{1.1}$$

$$\mathit{Some}(C, A) \tag{1.2}$$

However the inference rule taking the above premises and 1.1 as a conclusion is logically correct, whereas an inference rule taking the above premises and 1.2 as a conclusion is, due to the modern view, correct only with the additional assumption that the terms used in the premises are non-empty.

The reason for the above is the modern and commonly accepted logical representation of the categorial sentences in predicate logic. Thus the Aristotelian copula "all" is understood as a $\langle 1, 1\rangle$ generalized universal quantifier "all" and "some" as a $\langle 1, 1\rangle$ generalized existential quantifier "some". Generalized quantifiers are defined in terms of classes of models (so-called "Lindström definition" (Lindström, 1966)). In this style:

Definition 8. *A $\langle 1, 1\rangle$ generalized quantifier Q is a class of models over a vocabulary $\tau = \{P_1, P_2\}$ s.t. P_i, for $i \in \{1, 2\}$ is an unary predicate and Q is closed under isomorphisms, i.e. if M and M' are isomorphic τ-models, then $M \in Q \iff M' \in Q$.*

Since we consider only categorial sentences, we only apply quantifiers to atomic formulas. Suppose that ϕ_1, ϕ_2 are first order atomic formulas. Semantics for $\langle 1, 1\rangle$ generalized quantifiers is given in the following way:

$$M \models Qxy\phi_1(x)\phi_2(y) \iff Q_M(\phi_1^M(x), \phi_2^M(y)), \tag{1.3}$$

[5] In the case of one-premise arguments we have 4 (out of possible 16) inferences valid with the assumption that domains are non-empty, only 2 of them correct without any restrictions.

where M is a model and

$$\phi_i^M(x) = \{a \in M | (M, a) \models \phi_i(x)\}$$

The generalized quantifiers "all" and "some" are then defined as the following classes of models:

Definition 9. $All = \{(M, A, B) : A, B \subseteq M, A \subseteq B\}$

Definition 10. $Some = \{(M, A, B) : A, B \subseteq M, A \cap B \neq \emptyset\}$

Now if Aristotelian copulae are understood as generalized quantifiers, then a categorial sentence represented as $Q(A, B)$ is equivalent to $Qxy(A(x), B(y))$, (where Q stands for a generalized quantifier).

The logical difference between the generalized quantifiers "some" and "all" on the one hand and the existential quantifier $\exists$ or the universal quantifier $\forall$ on the other hand should be noted. The first are of type $\langle 1, 1\rangle$ and the second of type $\langle 1\rangle$, which means that the first applies to two formulas and binds one variable in each, and the second binds one variable in one formula. However "all" and "some" are definable in first order logic. Hence the following equivalences hold:

$$All(A, B) \iff \forall x(A(x) \rightarrow B(x)) \tag{1.4}$$

$$Some(A, B) \iff \exists x(A(x) \wedge B(x)) \tag{1.5}$$

$$No(A, B) \iff \neg\exists x(A(x) \wedge B(x)) \tag{1.6}$$

$$Somenot(A, B) \iff \exists x(A(x) \wedge \neg B(x)) \tag{1.7}$$

Now, both universal categorial sentences $All(A, B)$ and $No(A, B)$ are true and both particular sentences are false in empty domains. Therefore in the modern, *Revised Square of Opposition* only contradictories are preserved, whereas all the other dependencies are lost – including of course subalternations, which is why all the syllogisms based on the inferences of $Some(A, B)$ from $All(A, B)$ and $Somenot(A, B)$ from $No(A, B)$ are invalid.

Figure 1.2 pictures the Revised Square of Opposition.

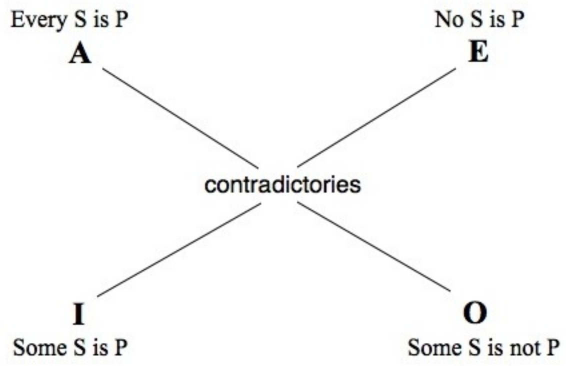


Figure 1.2: The Revised Square of Opposition

1.2 Categorial sentences with quantifier "most"

1.2.1 Semantics

Our aim is now to extend the language of Aristotelian syllogistic by adding the quantifier "most". We would like to be able to reason using the following categorial sentences:

1. "Most A's are B."
2. "Most A's are not B."

Such sentences cannot be expressed in first order logic. As a generalized quantifier "most" is like Aristotelian quantifiers of type $\langle 1, 1\rangle$ defined as follows.

Definition 11.

$$Most = \{(M, A, B) : A \subseteq M, B \subseteq M, (|A \cap B| > |A - B|)\}$$

Now Let M be a model and A, B predicates, then:

$$M \vDash Most(A, B) \iff (|A^M \cap B^M| > |A^M - B^M|) \tag{1.8}$$

What is characteristic, people indeed seem to use this quantifier in the above meaning. They rarely have any problems with determining, for example in a case of making a decision by voting, whether it was really "most" that voted in a particular way or not. However, asked how many "most" is, they begin to have a real problem. One can hear for example that it is $50\% + 1$. That such a definition is false can be easily shown. Let us take a set of five persons. $50\% \cdot 5 + 1 = 3.5$, whereas for the sentence: *"Most people (of five) voted for the proposal"* to be true it is enough that three persons voted for the proposal. Certainly there is no need for slicing anyone into two halves.

Some problems connected with Definition 11 should be mentioned. The first is that it is problematic in infinite domains. Suppose that A is an infinite set, e.g. the set of natural numbers, and B is a predicate "prime". Then a sentence "Most natural numbers are not prime" seems to be intuitively true, whereas in the sense of cardinalities it is false. The similar controversies arise in infinite domains also for other quantifiers, for instance "more" (for analogous example as for "most" see (Szymanik, 2009)), or "twice as many". "There are twice as many integers as natural numbers (+1)" seems true in the intuitive sense, but is false in the sense of cardinalities. Therefore we restrict our considerations to finite universes, which – as far as natural language is concerned – does not seem to be a controversial assumption. The reason for this is that when refereing to objects from the surrounding world we usually refer to finite sets, or at least finite (and thus conceivable) representations of infinite sets.

The other problems are of a more pragmatic nature. The above definition presupposes that the domain (A) is dichotomously divided into B and non-B. The question is what happens when we have a domain A, which is logically divided (which means that the division in exclusive and exhaustive) into more than 2 subsets. Let us suppose that we have for example a 10-element domain A (10 people at the conference) of which 4 are German, 3 are Dutch, 2 are Polish and 1 is French. It seems pragmatically totally reasonable to say that *most participants of the conference are German* (or maybe that *there are Germans in majority at a conference*).[6] Whether this is really the case in the natural reasoning should be empirically checked in further research. Still, the interpretation of "most" in terms of the division of a set appears to be an interesting option to consider.[7]

The other pragmatic gap of the above definition is the scalar implicature of the quantifier "most", i.e. the fact that in natural language "most" implicates or even means "not all". A detailed explanation of this issue will be given in part 1.3.

1.2.2 Negation

Since syllogistic is closed under negation we would like to consider also negative sentences with the quantifier $Most$.[8] The question is: which form of such sentences should we choose?

There are three important forms of negation in natural (and logical) languages (Peters, Westerstahl, 2006) and they are depicted in what is called a *Modern*

[6] I would like to thank to Henk Zeevat for this observation.

[7] Probably this could be also just ambiguity of the world "most", and hence such utterances would rather mean "The largest number was that of Germans".

[8] We will use $Most$ for denoting "most" defined as in Def. 11.

Square of Opposition for quantifiers: outer negation, inner negation and dual quantifiers.

Definition 12. *Let Q be a generalized quantifier of type $< 1, 1 >$.*
An outer negation of Q is a quantifier $\neg Q$, s.t.
$\neg Q(A, B) \iff \neg[Q(A, B)]$
An inner negation of Q is a quantifier $Q\neg$, s.t.
$Q\neg(A, B) \iff Q(A, M - B) \iff Q(A, \bar{B})$
A dual quantifier of Q is a quantifier Q^d, s.t.
$Q^d(A, B) \iff \neg(Q\neg)(A, B) \iff (\neg Q)\neg(A, B)$.

A quantifier together with its all three negations creates a (modern) "square of opposition". Thus outer negations lie in the square on diagonals, inner negations horizontally on opposite sides and duals vertically. For example the square for *All* is {*All*, *notAll*, *Some*(All^d), *No* ("All not")}

Definition 13. *Assume Q is a generalized quantifier of type $< 1, 1 >$. Then the (Modern) Square of Opposition for Q is the set $\{Q, \neg Q, Q\neg, Q^d\}$.*

The Square of *Most*

Let ϕ be a sentence saying: "Most A's are B-s". To express a contradictory sentence we could negate ϕ in two ways: by putting negation in front of the sentence (i.e. negating the whole sentence "directly") or by using a dual quantifier (and one more inner negation). The first is in natural language usually done by an expression "it is not the case that", or "it is not true that". As a result we obtain e.g. ϕ': "It is not the case that most A's are B." Such sentences are however known as "abnormal negations"[9] and presumably rarely appear in natural language. The equivalent of the above with "normal negation" would be a sentence with an outer negation of *Most*, i.e. "Less than half or half A's are B." This seems however to be quite a complex sentence, because "less than half or half" is a disjunction of two simple quantifiers.

"Normal negations" of sentences are also obtained with the use of dual quantifiers. Thus "normal negation" of ϕ would be here ϕ'': "$Most^d$ A's are not B", where $Most^d$ is a dual quantifier for *Most*. The problem with the last is that there is no "simple" natural language quantifier that is dual for *Most*. This is again a disjunction of quantifiers, i.e. "more than half or half". ("More than half or half A's are not B-s".) Thus neither ϕ' nor ϕ'' seems to be a good candidate for a negative sentence with the quantifier "Most" to be added to our syllogistic.

Our proposal is that this role should be played by a sentence ψ: "Most A's are not B", which uses the inner negation of *Most*. Of course the meaning of ψ is different from the meaning of ϕ' and holds if and only if $(Most\neg)(A, B)$. We further use a quantifier *Mostnot* for denoting $Most\neg$. We preserve the negative

[9]The distinction between normal and abnormal negations was introduced by Barwise (1979).

form of this sentence (i.e. "Most A's are not B"), although there is a "positive" analogue for *Mostnot* in natural language, namely a quantifier "minority" or "less than half". The sentence " Most A's are not B" is equivalent to the sentence "Less than half of A's are B" . We want however to preserve this negative meaning of this sentence for some reasons, one of them being the fact that negation seems to have an outstanding influence on the cognitive processes of language comprehension, the other that what we are investigating in our research is the natural-language quantifier "most" and not quantifiers "more than half" or "less than half". The latter is an important point since, although "most" is defined as "more than half", it does not necessarily mean that these quantifiers are understood in natural language or behave in reasoning in exactly the same way. For instance many of the pragmatical characteristics of "most" probably do not apply to "more than half" – e.g. the scalar implicature.

Let us define explicitly semantics for *Mostnot*. $Mostnot(A, B)$ we read as "Most A's are not B".

Definition 14. *Let M be a model and A, B – predicates. Then*
$M \vDash Mostnot(A, B) \iff (|A^M \cap B^M| \leq |A^M - B^M|)$.

Hence the main difference between the above and the outer negation of $Most(A, B)$ is that in the case of inner negation the relation is strict and in the case of outer negation it is not.

Now we may define the square for *Most*:

$Square(Most) = \{Most, \neg Most$ (less than half or half), $Most\neg = Mostnot$ (less than half), $Most^d$ (more than half or half)$\}$

1.2.3 Monotonicity

Monotonicity is treated as one of the most important properties of natural-language quantifiers. As syllogistic is sometimes named "monotonicity with addition of an existential import", it is worth to pay attention to this special property of *Most*.

Definition 15. *Let Q be a $< 1, 1 >$ generalized quantifier. Then:*

1. *Q is upward monotone in the first argument ($\uparrow Q$) if and only if: if $Q_M(A, B)$ and $A \subseteq A' \subseteq M$, then $Q_M(A', B)$ (persistence).*
2. *Q is downward monotone in the first argument ($\downarrow Q$) if and only if: if $Q_M(A, B)$ and $A' \subseteq A \subseteq M$, then $Q_M(A', B)$ (anti-persistence).*
3. *Q is upward monotone in the second argument ($Q \uparrow$) if and only if: if $Q_M(A, B)$ and $B \subseteq B' \subseteq M$, then $Q_M(A, B')$.*
4. *Q is downward monotone in the second argument ($Q \downarrow$) if and only if: if $Q_M(A, B)$ and $B' \subseteq B \subseteq M$, then $Q_M(A, B')$.*

One can quite easily observe that $Most$ is not monotone with respect to the first argument. Thus it is neither persistent nor anti-persistent. This can be checked using some syllogistic patterns. Let us take the following set of premises:

$$Most(B, A)$$
$$All(C, B)$$

Although the minor premise says that $C \subseteq B$, it does not follow that $Most(C, A)$. Thus $Most$ is not downward monotone in the first argument. If we change the argument, so that we take a superset of B, changing the minor premise to $All(B, C)$, we still cannot conclude $Most(C, A)$. Thus $Most$ is not upward monotone in the first argument. However, $Most$ is upward monotone in the second argument, namely if $Most(B, A)$ and $A \subseteq C$ (i.e. $All(A, C)$), then it follows that $Most(B, C)$). Similarly, the quantifier $Mostnot$ is non-monotone in the first argument and is downward monotone in the second, namely if $Mostnot(B, A)$ and $C \subseteq A$, then $Mostnot(B, C)$. What follows $\neg Most$ ("less than half or half") is downward monotone in its second argument (and not monotone in the first), whereas M^d ("more than half or half") is upward monotone in its second argument (and not monotone in the first).

Summarizing, monotonicity provides us with inferences about superdomains or subdomains, namely:

1. $Most(A, B)$ and $B \subseteq C$, then $Most(A, C)$
2. $Mostnot(A, B)$ and $C \subseteq B$, then $Mostnot(A, C)$

1.2.4 Existential meaning

The last characteristic of $Most$ to which we pay attention is that sentences with this quantifier are false in empty domains. Given Definitions 11 and 14, if at least one of the sets (A or B) is empty, then the sentences "Most A's are B" and "Most A's are not B" are always false. Hence also the following inferences hold:

$$\frac{Most(A, B)}{Some(A, B)} \qquad \frac{Mostnot(A, B)}{Somenot(A, B)}$$

The above is an important observation since in the experimental part we will be comparing the frequency with which subjects infer "some" from "most" with the frequency with which they infer "some" from "all" – whilst the latter inference is, according to the canonical view, invalid without the additional assumption of non-empty domains.

1.3 Scalar implicature

Scalar implicature is a *quantity implicature* (so one of the *conversational implicatures*) and is derived from one of the Gricean *conversational maxims*, namely the *Maxim of Quantity:*

> *Be as informative as required.* (Grice, 1989)

In other words the maxim of quantity is a commandment to make our contribution to the conversation as informative as required. In effect it involves the fact that in natural communication we usually assume that we are given *all* the relevant (the *Maxim of Relevance*) information.

Scalar implicature is based on the use of an informationally weak term in an *implicational scale* (a set of lexical items of the same category ordered according to their informativeness) (Levinson, 1983). For example, in the case of quantifiers the following scale is usually given ⟨all, most, many, some⟩ (from the strongest to the weakest). The use of the weak term *implicates* (i.e. implies as an implicature) that all similar utterances using an informationally stronger term are not true since, according to the Maxim of Quantity, a speaker would be required to make a stronger, more informative utterance if a true one were available. Thus if one says: (1) "Some girls wore skirts", the quantifier "some" in this sentence implicates that *not all* girls wore skirts. Similarly (2) "Most girls wore jeans" implicates that *not all* girls wore jeans, as both "some" and "most" are lower on the implicational scale than "all". What is more, (1) should implicate also that "Most girls did not wear skirts", since "some" is weaker than "most". Throughout we call "some not" the *weak implicature* of "some", and "most not" the *strong implicature.* Analogously for "some" and "most" – implicatures of "some not". Thus "weak" refers to implicating a sentence with the weaker item on the implicational scale (namely "some") and "strong" to implicating a sentence with the stronger item ("most"), so a sentence which is stronger in a logical sense.

The mentioned entailments are not valid in the classical meaning of a correct inference. Given Definition 11 and 10, both sentences "Most A's are B" and "Some A's are B" are true if all A's are B, and thus do not imply that $\neg All(A, B)$. On the other hand, the stronger expressions are supposed to imply all the weaker ones on the scale. Thus the following inferences hold (the inference relation is transitive):

$$\cfrac{\cfrac{All(B,A)}{Most(B,A)}}{Some(B,A)}$$

What is worth noting is that both inferences $All/Some$ and $All/Most$[10] are dependent on the existential import of "all". If we agree that "all" contrary to

[10]This notation should be understood as an inference rule with a premise of one form and a conclusion of another, e.g. $All/Some$ translated to natural language yields: $\frac{\text{"All A's are B"}}{\text{"Some A's are B"}}$

"most" and "some" is true in empty domains, then neither of these inferences is valid.

Similarly we may consider the negative quantified sentences, which would create the following implicational scale ⟨"some..not", "most..not", "no"⟩. Assuming Definition 14, "Most A's are not B" is true also when $A - B = \emptyset$ (A,B non-empty), similarly for $Somenot(A, B)$, thus in a classical sense neither $Mostnot(A, B)$ nor $Somenot(A, B)$ implies $\neg No(A, B)$. On the other hand, the following inferences hold:

$$\frac{\dfrac{No(B, A)}{Mostnot(B, A)}}{Somenot(A, B)}$$

Again both $No/Mostnot$ and $No/Somenot$ are dependent on the solution of the "existential import" problem.

The crucial property of implicature that is usually assumed is its *non-monotonicity.* In non-monotonic reasoning an inference may be affected by adding new premises. In effect a conclusion may not follow from an extended set of premises, although it followed from the original one. Let us see how it works in context.

> *A: Most girls wore dresses.*
>
> *B: Most? Well, I suspect all, I heard that it was the dressing code requirement for women to wear a dress for this meeting.*
>
> *A: Ok, I didn't know about this, so maybe even all of them, I was not paying that much attention to people's clothes.*

In this short conversation, the interlocutor B implicates from A's sentence that not all girls wore dresses. As it contradicts her knowledge that there was a formal requirement for women to wear a dress, she immediately mentions this fact. A is however not confused and admits that there may be a gap in her knowledge, and thus that the universal sentence is still possible. Thus A both agrees with the consequence of what she says which is suggested by B (hence she agrees with the implicature of "most") and also she can easily extend her database so that the initial conclusion is rejected but a new one is inferred. The question is whether we want to claim that the first sentence given by A is false (since it implicates that not all girls wore dresses and actually all of them did) or if we agree that this apparent contradiction is not a real one, but just an effect of non-monotonicity at work.

Due to their non-monotonic character, implicatures are usually formalized with the use of logic programming. Adding implicature as a definitional condition to the semantics of a given expression is usually considered a false attitude. In the case of the quantifier implicational scale, it would result in losing all the entailments where a weaker expression is entailed from the stronger one. Namely if we defined "some" and "most" so that they imply "not all", we could not infer categorial sentences with these quantifiers from universal categorial sentences – even if the additional assumption that domains are non-empty were made. Since it is not consistent with our language intuitions, we are unwilling to agree to such a solution. What is more, such quantifiers would not be monotone (see 1.2.3). More precisely, let $Most'$, $Mostnot'$, $Some'$ and $Somenot'$ be defined as follows:

Definition 16. *(Alternative "most" and "some")*

$$Most'(A,B) \iff [(|A \cap B| > |A - B|) \wedge \neg(A \cap B = A)]$$
$$Mostnot'(A,B) \iff [(|A \cap B| < |A - B|) \wedge (A \cap B \neq \emptyset)]$$
$$Some'(A,B) \iff [(A \cap B \neq \emptyset) \wedge \neg(A \cap B = A)]$$
$$Somenot'(A,B) \iff [(A - B \neq \emptyset) \wedge (A \cap B \neq \emptyset)]$$

It is clear, that neither $All/Most'$ nor $All/Some'$ (and analogously $No/Mostnot'$ nor $No/Somenot'$) are valid inferences. We lose hence monotonicity. $All(B,C), Most'(A,B)/Most'(C,A)$ is an invalid inference, whereas it would be valid if instead of $Most'$ we used $Most$. The same for $All(C,B), Mostnot'(A,B)/Mostnot'(A,C)$. Similarly, it works for syllogisms with $Some'$ (resp. $Somenot'$). Whereas $Some$ is upward monotone in both arguments ($Somenot$ – downward monotone in both arguments), $Some'$ and $Somenot'$ are not monotone in their second arguments.

With such an approach, we would lose many syllogisms with "most", which seem quite consistent with natural-language intuitions. (See section 1.5.) Later we will try to establish to which extent the scalar implicatures may be embedded in the semantics and what kind of formal modeling seems the best for this reasoning phenomenon.

In our experiment we check implicature "not all" for both "most" and "some" together with the negative counterpart, namely implicature "not no" for "most not" and "some not". Furthermore, we check also the strong scalar implicature of "some". Consequently the following direct quasi-inferences, which are not valid in a classical sense, but can be justified on a basis of pragmatic reasons, are considered in our study:

$$\frac{Most(A,B)}{\begin{array}{c}Somenot(A,B)\\(\neg All(A,B))\end{array}} \qquad \frac{Mostnot(A,B)}{\begin{array}{c}Some(A,B)\\(\neg No(A,B))\end{array}}$$

$$\frac{Some(A,B)}{\begin{array}{c}Somenot(A,B)\\(\neg All(A,B))\end{array}} \qquad \frac{Somenot(A,B)}{\begin{array}{c}Some(A,B)\\(\neg No(A,B))\end{array}}$$

$$\frac{Some(A,B)}{Mostnot(A,B)} \qquad \frac{Somenot(A,B)}{Most(A,B)}$$

1.4 Existential import

Existential import is an ambiguous notion. Traditionally it refers to the question whether the universal categorial sentence "All A's are B" implies that A's exist. The origin of this question lies in the "subalternation" relation of the Aristotle's Square of Opposition. More precisely the question is dependent on the following inference, later named in the thesis "so-called existential import":

$$\frac{All(A, B)}{Some(A, B)}$$

The argumentation proceeds as follows. Since "some" is existentially loaded, then also "all" has to be, because otherwise we would lose a part of the Square of Opposition, i.e. we would loose all relations in the Square but contradictories. (See section 1.1.2 and figure 1.2.)

Actually in the presence of empty terms, even if existential import is granted to the A-sentence ("All A's are B"), the whole square still requires revision. Let us suppose that there are no A-s. Then the I-sentence ("Some A's are B") is false. What follows the contradictory sentence E ("No A's are B") is true and hence its subaltern O ("Some A's are not B") must be true, but it cannot be, since there are no A-s.

The above contradiction, however, arises only assuming the modern interpretation of categorial sentences, in which the O-sentence gets, similarly to the I-sentence, an existential reading. The Aristotelian doctrine of the Square was completely coherent in the presence of empty terms. This is because, according to the Aristotle's interpretation, the O-form lacked existential import and was true in empty domains. Aristotle's articulation of the O-sentence was not the familiar "Some A's are not B"; it was rather "Not every A is B" – and thus was understood as universal. Then the E-sentence "No A-a are B" – true in empty domains – entailed the O-sentence in its universal formulation. Since "All A's are B" (or rather: "Every A is B" – in the Aristotle's formulation) was granted existential import, and thus was false if A was empty, the consistency of the Square was preserved.

It is worth noting that in the case of empty domains the Aristotelian Square is inconsistent also if all the categorial sentences are granted existential import. Let us assume that there are no A's. Then all four sentences are false, and hence although subalternation and contraries are preserved, contradictories and subcontraries are not.

However, what does it mean that a sentence is *granted existential import*? In the wide, classical sense, it means only that this sentence in that or another way implies non-emptiness of its domain.

Definition 17. *Let ϕ be a categorial sentence, so $\phi := Q(A, B)$. Then, ϕ is characterized with* existential import *(or is* existentially loaded*) if and only if ϕ implies that there are A's.*

Such an definition is, however, rather vague and allows different interpretations and hence also different solutions of the existential import problem.

1. We can agree for the Aristotelian solution, namely that only the A-form and I-form imply that there are A's, whereas E- and O-forms lack existential import.

   There are a few gaps in this solution. First of all we may ask, what kind of logical form does an O-sentence have, if it is supposed to be universal. In classical predicate logic a sentence "Not every A is B" obtains a form $(F)\,\neg\forall x(A(x) \to B(x))$, but then it is equivalent to $(F')\,\exists x(A(x) \wedge \neg B(x))$. However, one could try to interpret negation in a different, non-classical, way, so that (F) is read as e.g. "It is absurd to assume that every A is B" and is not equivalent to (F'). But how can it be contradictory to "All A's are B"? If we negate F, can a negation interpreted in this way be eliminated by a double negation elimination rule?

   Furthermore, one and the most straightforward interpretation of existential import would be to assume that it has a form of an explicit condition $\exists x A(x)$ occurring in the logical form of a categorial sentence. Then, if in the spirit of the Aristotelian solution, we grant existential import only to the positive categorial sentences and refuse it to the negative, we get the following readings:

   $$\mathit{All}(A, B) \iff A \subseteq B \wedge \exists x A(x) \tag{1.9}$$

   $$\mathit{Some}(A, B) \iff \exists x(A(x) \wedge B(x)) \tag{1.10}$$

   One might doubt whether indeed 1.9 is what we mean when saying "All A's are B", but such an approach leads to even more complications. Now to preserve the whole Square of Opposition the sentences: "No A's are B" and "Some A's are not B" get the non-intuitive disjunctive forms:

   $$\mathit{No}(A, B) \iff \neg\exists x(A(x) \wedge B(x)) \vee \neg\exists x A(x) \tag{1.11}$$

   $$\mathit{Somenot}(A, B) \iff \exists x(A(x) \wedge \neg B(x)) \vee \neg\exists x A(x) \tag{1.12}$$

   Now all the relations in the Square are preserved. The problem is that 1.11 and 1.12 seem rather inadequate as expressions of logical forms. It seems contr-intuitive to say that we are stating in those cases additionally a disjunctive component: "or there are no A's".

2. We can agree with the modern reading of categorial sentences (see section 1.1.2) and thus grant existential import to both particular sentences and refuse it to the universal ones. Then we are left with the Revised Square of Opposition in which only contradictories are preserved and the inferences *All/Some* and *No/Somenot* are invalid.

   The problem lies mainly in the fact that this view does not take into account the natural-language intuitions. It seems that for average users of language there is no obvious difference with respect to emptiness of domains between sentences with "all" and sentences with "some". The only real difference between these quantifiers is of a scalar nature – namely that "all" refers to the whole domain and that "some" refers to a subset of it. Moreover this intuition is consistent with the assumption of how scalar implicatures of these quantifiers are supposed to work. On the implicational scale stronger items are supposed to entail weaker and hence "all" should entail "some".

   Furthermore, why should there be any truth-value difference between the two sentences "All round squares are squares" and "Some round squares are squares"? Intuitively, they are either both true (analytically!) or they are both false since they both refer to non-existent objects (or both deprived of truth value). Later we will see that there is no real reason to refuse truth to sentences of form "Some A's are B" just because A's are non-existent objects.

3. We can decide that all terms in natural language are non-empty, so that *for every term t, there is x s.t. it is true that $t(x)$.* Here we come to the key issue – namely which are the *empty terms*. More space will be devoted to this problem also later in section 2.2.2, now just a short introduction.

   Roughly speaking, by an "empty term" we mean a term without reference, i.e. such that there are no entities to which it applies. Among most famous philosophical examples of empty terms one may consider a common noun "unicorn" for which alleged denotation some ontological realists are supposed to constantly and desperately search but have no hope for success, unless they can use for a bait a virgin. What is silently assumed is that empty terms have no denotation in the real (physical) world, thus the ontological issue is here implicitly solved. Only physical objects exist and are *designata* of terms, precisely of non-empty terms. Thus, terms naming fictitious (e.g. elves) or contradictory (e.g. round squares) objects are empty (whatever "naming" means in this context) and do not refer to anything. Therefore, is is assumed that existential sentences with the use of such terms are false and this is how the whole problem of existential import arises.

   In this way our ontology decides about our logics and existential import turns out to be what is commonly known as ontological commitment. However, if "elves", "Zeus", "unicorn" are empty terms, what do we refer to while uttering (obviously true) sentences like "elves do not exist"? And also – what are we talking about in mythological or fictional discourse?

The easy and naive solution would be to agree that *all terms are non-empty*. But in this context it would just mean that everything exists – that there is no ontological difference between elves and humans, unicorns and horses, dragons and reptiles. Most people, however, can see a great ontological difference between these objects and those who have problems with distinguishing reality from fiction, humans from elves or fear that dragons could come to their houses and eat them, usually require a psychiatric treatment. Non-empty and empty terms differ essentially in the ontological sense, even though one could agree that the latter refer to *something* – just something non-existent.

4. Finally we could move the burden of the problem from the question whether "all" should be granted existential import to the question whether "some" is existentially (or ontologically) loaded. Then, we could try to treat quantification as ontologically neutral, although we would have to face again the problem of ontological status of possible (or non-existent) objects and also answer the question how to interpret the quantifier "some".

   If "some" is understood as a $\langle 1, 1\rangle$ generalized quantifiers defined as in 10, then indeed by $Some(A, B)$ we state *non-emptiness of some set* – namely of the intersection of A and B, which are subsets of the universe of some model. In this sense this quantifier is indeed existentially loaded. However, we may always ask what kind of model we presume or *presuppose* here. And why the physical world in which we live! Then also the problem of existential import can be reformulated. What we could ask is, whether we really should understand existential import of a given sentence in the sense of Definition 17, so that the thesis that some entities exists (implicitly: in the real world) can be inferred from this sentence. What other form can this question get we will try to show in Chapter 4.

1.5 Syllogistic with "most"

1.5.1 Classically valid syllogisms with $Most$ and $Mostnot$

There are very many possible syllogisms with $Most$ and $Mostnot$ that are syntactically correct. We have 6^3 possible choices of quantifiers in premises and conclusion, so combinations of form $Q_1, Q_2|Q_3$, where $Q_i \in \{$All, Some, Somenot, No, Mostnot, Most$\}$, and 4 syllogistic figures, which gives us $6^3 * 4 = 864$ possible syllogisms. However only 21 are semantically valid inferences, namely true in all possible models (with an assumption that domains are non-empty). In fact only: 16 two-premises syllogisms (and 2 direct inferences[11]) with quantifiers *Most* and *Mostnot* are correct in an "absolute sense", namely independently of any additional assumptions about domains.[12]

The candidate syllogisms can be generated by a simple computer program, and next checked by hand. The program has definitions of quantifiers implemented and for each possible syllogism it checks whether there exist a model of cardinality N where premises are true but a conclusion is not. This is done by checking all possible models where A, B, C - non-empty subsets of $M = \{1...N\}$.[13] Then if answer is "yes", the syllogism cannot be valid. It turns out that the program gives correct output already for $N = 3$ [14] The 21 syllogisms generated by such a program have indeed been checked to be valid inferences. Thus we can conclude that the following holds:

Conjecture 1. *Let S be a syllogism in a language with Most, over predicates $A, B, C \neq \emptyset$. Whenever S is valid in all models of cardinality ≤ 3, then it is valid in any model.*[15]

Below we list all valid syllogisms with *Most* and *Mostnot*. (24 Aristotelian syllogisms are omitted.) First we consider one-premise direct inference, and subsequently canonical two-premise syllogisms categorizing them according to: (a) occurrence of *Most* or *Mostnot* in premises or a conclusion (b) figures.

Direct inferences

We write in bold those conclusions that can be inferred only with the assumption that domains are non-empty.

$$\frac{All(A,B)}{\mathbf{Most(A,B)}} \qquad \frac{No(A,B)}{\mathbf{Mostnot(A,B)}}$$

$$\frac{Most(A,B)}{Some(A,B)} \qquad \frac{Mostnot(A,B)}{Somenot(A,B)}$$

[11] Out of 36 syntactically possible and 4 semantically valid with the assumption of non-emptiness of domains

[12] Considering valid inferences with *Most* we do not take into account purely Aristotelian syllogisms any more.

[13] Since there are $2^N - 1$ non-empty subsets of M and we try all possible subsets for A, B, C, there are in general $(2^N - 1)^3$ models of size N.

[14] Such a program checks validity for $(2^3 - 1)^3 = 343$ models

[15] We do not have a mathematical argument for this to hold right now, but we checked by hand that all the generated syllogisms are indeed valid inferences – namely valid also for domains of cardinality > 3. Additionally, we checked also that the program generates exactly the same output for $N = 7$, and it can be shown that 7-element models are big enough for checking validity of traditional syllogistic. The main idea of the proof is that we are interested only in (non-)emptiness of each part of a Venn Diagram and there are only 7 of them.

Two-premise syllogisms with the quantifiers $Most$ **and** $Mostnot$

1. Syllogisms with one Aristotelian quantifier and $Most$ in premises:

   - I Figure

$$\begin{array}{c} All(B,A) \\ Most(C,B) \\ \hline Most(C,A) \\ Some(C,A) \end{array} \qquad \begin{array}{c} No(B,A) \\ Most(C,B) \\ \hline Mostnot(C,A) \\ Somenot(C,A) \end{array}$$

   - II Figure

$$\begin{array}{c} No(A,B) \\ Most(C,B) \\ \hline Mostnot(C,A) \\ Somenot(C,A) \end{array}$$

   - III Figure

$$\begin{array}{c} All(B,A) \\ Most(B,C) \\ \hline Some(C,A) \end{array} \qquad \begin{array}{c} Most(B,A) \\ All(B,C) \\ \hline Some(C,A) \end{array} \qquad \begin{array}{c} No(B,A) \\ Most(B,C) \\ \hline Somenot(C,A) \end{array}$$

   - IV Figure

$$\begin{array}{c} Most(A,B) \\ All(B,C) \\ \hline Some(C,A) \end{array} \qquad \begin{array}{c} No(A,B) \\ Most(B,C) \\ \hline Somenot(C,A) \end{array}$$

2. Syllogisms with one Aristotelian quantifier and $Mostnot$ in premises:

   - II Figure

$$\begin{array}{c} All(A,B) \\ Mostnot(C,B) \\ \hline Mostnot(C,A) \\ Somenot(C,A) \end{array}$$

   - III Figure

$$\begin{array}{c} Mostnot(B,A) \\ All(B,C) \\ \hline Somenot(C,A) \end{array}$$

3. Syllogisms with $Most$ or $Mostnot$ in both premises:

   - III Figure

Rescher's syllogism	**"negative" Rescher's syl.**
$\begin{array}{c} Most(B,A) \\ Most(B,C) \\ \hline Some(C,A) \end{array}$	$\begin{array}{c} Mostnot(B,A) \\ Most(B,C) \\ \hline Somenot(C,A) \end{array}$

4. Syllogisms with Aristotelian premises and a *Most* or *Mostnot* in the conclusion

   - I Figure

$$\frac{\begin{array}{c} All(B,A) \\ All(C,B) \end{array}}{\textbf{Most (C,A)}} \qquad \frac{\begin{array}{c} No(B,A) \\ All(C,B) \end{array}}{\textbf{Mostnot(C,A)}}$$

   - II Figure

$$\frac{\begin{array}{c} All(A,B) \\ No(C,B) \end{array}}{\textbf{Mostnot(C,A)}} \qquad \frac{\begin{array}{c} No(A,B) \\ All(C,B) \end{array}}{\textbf{Mostnot(C,A)}}$$

   - IV Figure

$$\frac{\begin{array}{c} All(A,B) \\ No(B,C) \end{array}}{\textbf{Mostnot(C,A)}}$$

1.5.2 Additional inferences depending on implicature

The above syllogistic does not take into account scalar implicature that was considered in section 1.3. Taking implicature into account we will get 10 more quasi-inferences based on the assumption that "most" implicates "not all" and that "most not" implicates "not exclusive". In what follows, we mark with * those conclusions that dependent on the implicature of the quantifier *Most* in the premise. For instance $All(A,B), Mostnot(C,B)/Some(C,A)$ is not a valid inference pattern, whereas $All(A,B), Mostnot(C,B)/Some^{*}(C,A)$ is pragmatically justified. (We first implicate $Some(C,B)$ from $Mostnot(C,B)$ and then from $All(B,A)$ and $Some(C,B)$ we can infer $Some(C,A)$.) As we can see all these inferences have some counterparts in Aristotelian syllogistic, e.g. the above-mentioned syllogism could be in a sense "reduced" to the following Aristotelian syllogism:

$$\frac{\begin{array}{c} All(B,C) \\ Some(C,B) \end{array}}{Some(C,A)}$$

Below we list the additional two-premise quasi-inferences based on the implicature:

1. First figure

$$\frac{\begin{array}{c} All(B,A) \\ Mostnot(C,B) \end{array}}{Some(C,A)^{*}} \qquad \frac{\begin{array}{c} No(B,A) \\ Mostnot(C,B \end{array}}{Somenot^{*}(C,A)}$$

2. Second Figure

$$\frac{\begin{array}{c}All(A,B)\\ Most(C,B)\end{array}}{Somenot^*(C,A)} \qquad \frac{\begin{array}{c}No(A,B)\\ Mostnot(C,B)\end{array}}{Somenot^*(C,A)}$$

3. Third Figure

$$\frac{\begin{array}{c}All(B,A)\\ Mostnot(B,C)\end{array}}{Some^*(C,A)} \qquad \frac{\begin{array}{c}No(B,A)\\ Mostnot(B,C)\end{array}}{Somenot^*(C,A)} \qquad \frac{\begin{array}{c}Most(B,A)\\ All(B,C)\end{array}}{Somenot^*(C,A)} \qquad \frac{\begin{array}{c}Mostnot(B,A)\\ All(B,C)\end{array}}{Some^*(C,A)}$$

4. Fourth Figure

$$\frac{\begin{array}{c}No(A,B)\\ Mostnot(B,C)\end{array}}{Somenot^*(C,A)} \qquad \frac{\begin{array}{c}Mostnot(A,B)\\ All(B,C)\end{array}}{Some^*(C,A)}$$

Let us stress that we use here standard quantifiers $Most$ and $Mostnot$ and not $Most'$ and $Mosnot'$. Hence, we do not lose any of the classical syllogisms listed in section 1.5.1, which remain valid. The "star"-index in the conclusion only indicates that it is an implicature-base inference.

1.6 Summary

We introduced the preliminaries necessary for our thesis. We provided the theoretical assumptions that served as the basis for our experimental work: the syllogistic with $Most$ – as an extension of the Aristotelian syllogistic, the notion of *scalar implicature* of quantifiers and the problem of *existential import*. At this point we will leave the area of logic and go into the wild world of empiria. The next two chapters will be devoted to the experimental work, the purpose of which was to see to which extent our theory can be applied to natural language and if it cannot be, how our principles should be reformulated.

Chapter 2

First experiment: active competence

2.1 Introduction

2.1.1 Logical competence

The notion of *logical competence* (*LC*) was first used by Macnamara (1986). It is used in cognitive psychology, analogously to the notion of *linguistic competence* introduced by Noam Chomsky, to name a specific cognitive mechanism enabling to complete single logical tasks, for example performing reasoning in natural language according to some logical inference patterns. We assume that:

1. The human mind is equipped with such a mechanism.

2. This mechanism of logical competence consists of two basic functions:

   (a) *Passive logical competence*: an ability to evaluate, on the basis of given premises, a given conclusion as a correct or incorrect inference.

   (b) *Active logical competence*: nn ability to generate, on the basis of given premises, a correct conclusion.

Few things require explanation here – especially what we mean by a correct inference. It might seem that the notion of logical competence presupposes an absolute notion of logical correctness. The traditional approach to logical competence follows this attitude and, putting the main stress on how logical competence is divergent from logical correctness, defines the later with respect to classical logic. Thus a reasoning instance is correct if it is a classically valid inference.

In the psychological literature, it is a widespread view that reasoning proceeds according to some pre-established logical laws and thus may be either correct of incorrect, depending on whether it stays in accordance with these laws or not. A list of logical mistakes most frequently made by people is usually given together with putative explanations as to why people tend to make these very mistakes (and not some others). The famous *atmosphere effect* is an example. This is a notion used to explain people's preference to infer conclusions of a specific form from premises of a specific form. For instance the lack of particular conclusions inferred in the case of syllogisms with both universal premises is explained by saying that universal premises create a "universal atmosphere" and in this way they suppress the particular conclusion. (Nęcka, 2006) Such a view ignores the fact that in the classical sense validity of an inference with a universal premise and a particular conclusion requires an additional assumption of non-empty domains. Thus it classifies correct – in the light of the classical logic – reasoning as incorrect. Moreover, it also tries to explain human behavior by a very vague, if not meaningless, notion of "atmosphere". Such partial and unsuccessful theories are the result of a false conjecture that the correctness of a given piece of reasoning is somehow decided by the laws of the classical logic and disobeying these laws is a malfunction of our logical competence mechanism.

The root of such a view on reasoning is the assumption that logic is "domain-independent", which means that what is a correct inference does not depend on a specific domain. In other words it is supposed that inferences are correct or incorrect independently of the domain in which one reasons. Logic indeed is domain-independent in a sense that each *chosen logical system* (classical logic, many-valued logic, etc.) is such. However the *logic of reasoning*, thus the set of schemata used by humans, is domain-*dependent*, which means that the valid schemata depend on the domain: different schemata from different logical models are applied to different domains. There is no absolute logical system that can be recognized as the "logic of mind", no universal pattern applicable to all different forms of natural reasoning. Specific domains require specific interpretations.

It has been shown (most of psychology handbooks mention this as a fact: (Sternberg, 1996), (Nęcka, 2006)) that reasoning is indeed dependent on numerous factors, such as the linguistic *context* and *content* of a given reasoning, subject's knowledge or ontological status of alleged entities. Nevertheless, it is still frequently claimed that these phenomena are *disturbances* of logical competence and evidence that humans do not cope well with logic, and thus that our logical competence may or may not function in accordance with logical correctness.

However, if logical competence is a mechanism analogous to linguistic competence, the notion of logical correctness requires reformulation. A claim that some systematic discordance of human reasoning with classical inference schemes proves imperfection of our logical competence mechanism would be similar to telling that systematic ungrammaticality (in a sense of a language phenomenon not coherent with the prescriptive grammar) occurring in some ethnic language is a proof of dysfunction of human linguistic competence – a view contemporarily considered as false and obsolete. What is grammatical or ungrammatical is defined by our linguistic competence, so by some module we have in our heads determined partly by the Universal Grammar and partly by parameters of specific ethnic languages. Thus, this is grammar that should fit our linguistic competence and not the other way round. Similarly, it is our **logical competence that defines what is a correct inference and hence there is no "divergence" between logical correctness and logical competence.**[1]

Michiel van Lambalgen and Keith Stenning (2008) claim that human reasoning consists of two stages: *reasoning towards an interpretation* (establishing the domain in which one reasons and its formal properties) and *reasoning from an interpretation* (following the formal laws that are implied by the fixed interpretation). We agree with this view and propose that the mechanism of logical competence applies to both stages. First of all, this is our competence that defines logical schemata for different domains of reasoning on the interpretation stage. Secondly, the competence enables to follow the rules as soon as the formal model is fixed (however unconscious may be this process). One should note, that even if we use in reasoning various domain-dependent schemata, we still have to be able to adhere to some logical laws – the ones that apply to a given domain. Otherwise one would have to agree that humans are totally irrational. Moreover, the schemata that are chosen at the stage of establishing an interpretation need to have some nice logical properties. What kind of properties is a vast question, which we will not answer now, however, intuitively, those schemata cannot be totally arbitrary. Last but not least, the logical competence mechanism must be in a sense *universal*, which means that similar domains involve similar schemata that are used by all humans. Whether or not this mechanism is universal in a sense that there exists a general theory (analogous to the X-bar theory in linguistics) that would universalize the "logic of mind" (so would imply all the schemata that are used in various domains) is another question and remains open.

The very last thing that needs explanation is the above-mentioned distinction between *passive* and *active* competence. One may doubt if these are really two separate functions of the mechanism, however it seems very plausible that there is indeed a difference between the two processes: passive evaluation of an inference and its active generation. Such a phenomenon has been observed also in linguistics – the skill of generating grammatical utterances differs from the skill of evaluation whether a given expression is grammatical or not.[2]

[1] On the other hand not every ungrammaticality is accepted, and also not every illogicality will be. In linguistics one differentiates between *competence* and *performance* – a competence defines grammar, but a specific use of competence may of course be incorrect because of various reasons, e.g. complicated multiply embedded constructions, processing difficulties, memory malfunctioning etc.; analogously in the case of logical competence – a big number of premises for example may be a hindering factor and a source of possible mistakes.

[2] Such a difference can be found also between language generation and language comprehension mechanisms, which – since they can be separately impaired – are shown to be two distinct cognitive functions. There are kinds of aphasia consisting in dysfunction of production of grammatical language (both spoken and written) with a properly functioning language comprehension. This is called "Broca's aphasia". In a related syndrome patients with Wernicke's aphasia produce fluent, grammatical, but nonsensical (without semantical sense) speech and display problems with understanding language. In a sense, passive logical competence is more

2.1.2 Goals of the experiment

As far as the distinction between active and passive logical competence is concerned, some preliminary research was done in this field by Bucholc (2004), (see also (Mostowski & Bucholc, 2005)), who investigated logical competence with respect to Aristotelian syllogisms with empty terms. The author's aim was to investigate existential import of the quantifier "all" with respect to both passive and active competence. She claims that people behave differently when they are asked to generate a conclusion and differently when they are asked to evaluate it. Precisely, they may be disposed to evaluate a given conclusion as a consequence of premises even though they would not generate it themselves. It concerns especially the case when particular conclusions are inferred from universal premises with the additional assumption of non-emptiness of domains. Bucholc assumed in her study that inferences of the form *All/Some* are invalid in empty domains. Therefore, she tested direct inferences as well as syllogistic reasoning (especially moods with universal premises and particular conclusions) with both empty and non-empty terms. Her purpose was to investigate if human logical competence functions with accordance to the above correctness assumption. Basing on her experimental results she claims that at least with respect to the above-mentioned inference the active competence is more consistent with the correctness than the passive competence, since people are not willing to infer "some" from "all", especially when empty terms are used, although they tend to evaluate such inferences as correct. Thus the author claims also that the use of empty terms has an impact on existential import.[3] No significant statistical analysis is, however, given to confirm this thesis.

Although we find the assumptions of the above study dubitable – especially the absolute notion of logical correctness defined in terms of classical logic, the general idea of testing existential import depending on the domain was the first inspiration for doing this research. In our experiment we were interested in the following:

1. The so-called existential import of the quantifier "most" (compared with the so-called existential import of the quantifier "all") – we use the notion "so-called existential import" to name both: inferences *All/Some* and *No/Somenot* as well as inferences *Most/Some* and *Mostnot/Somenot*. The phrase "so-called" is used since we actually doubt that such infer-

similar to language comprehension than to evaluation of grammaticality. Evaluation of grammaticality requires some meta-knowledge about language, whereas evaluation of correctness of an inference does not necessarily require that kind of meta-knowledge. That is, it can be tested without involving any theoretical notions. While testing logical competence, one usually checks subjects' readiness to accept some conclusions in the light of given facts / knowledge / sentences without using the notions of "logical correctness" or "valid inference".

[3] or rather on inferences: *All/Some* and *No/Somenot*, called in the present study "so-called existential import"

ences have anything to do with existential import itself, which we rather understand as a kind of ontological commitment of those quantifiers. Of course there is an essential difference between this kind of inference in the case of "all" and in the case of "most". In the latter case, it is a classically valid inference not requiring any assumption about non-emptiness of domains. However, since we are going to compare these inferences, for ease of exposition, we will use henceforth the same notion "so-called existential import" for both "most" and "all".

2. Scalar implicatures of "some" and "most".

3. Complexity of syllogisms with "most" – which of them are difficult and which are easy for subjects.

For all reasoning experiments we prepared both: tasks with empty and non-empty terms to check domain-dependence of the considered inferences. The first experiment (or rather set of experiments) was designed to check the research questions with respect to active competence. In the second experiment (Chapter 3) we tested our hypotheses with respect to passive competence.

The "active" experiment described in this chapter should be treated more as qualitative than quantitative research. The replies given by subjects in active generation tasks were often too far from standard answers expected by the experimenter and they had to be interpreted before they could be counted. So the qualitative evaluation was an essential stage before any quantitative analyses could be done. Nonetheless we tried to analyze all our results also with respect to statistical significance.

2.1.3 General hypotheses

The following general hypotheses followed from our theoretical presumptions and constituted the background for the detailed project of the first experiment.

Hypothesis I: "Most" has stronger so-called existential import than "all", so the number of particular conclusions inferred from "most" should be higher than the number of particular conclusions inferred from "all".

First of all, assuming classical definitions of "most" and "all" in terms of $\langle 1, 1\rangle$ generalized quantifiers, "most" has, contrary to "all" but similarly to "some", existential meaning. This may be one reason for the expected effect. In may, however, happen that other observations (e.g. the lack of difference between the empty terms group and the control group), will not allow us to connect this effect with existential commitment. Then, if the hypothesis is still sustained, we will need to find some other explanation – e.g. in terms of scalar properties of the quantifiers.

Hypothesis II: An influence of scalar implicatures on inferences.

We expect implicature-based inferences (*Most* → *Somenot*, *Mostnot* → *Some*, etc.) to appear significantly in subjects reasonings. They may be dominated by classical inferences in complex two-premise tasks, however they should occur clearly in the direct-inference part.

Hypothesis III: Given the assumption that "some", contrary to the universal quantifier ("all" resp. "no"), is false in empty domains, so-called existential import should be stronger in the group with non-empty terms than in the group with empty terms. This means that we would expect less frequent occurrences of particular sentences inferred from universal premises, if terms used in tasks are empty. On the other hand no between-group difference would be evidence that so-called existential import is not dependent on emptiness of terms.

Additionally, if "most" has like "some" existential meaning, it should not be inferred from "all" in the case of empty domains either.

The other factors that should be taken into account in the analyses:

- Occurrence of negation in premises or/and in a conclusion can display an impact on a number of correct inferences (negative correlation – a negation makes a task harder).

- Difference in difficulty between syllogistic figures, e.g. the first figure is usually considered as the easiest for people, whilst the third and fourth figures are the most difficult.

2.2 Experimental layout

Three parts of the experiment were planned:

1. Picture task checking basic understanding of "most" in small domains.

2. Direct inferences.

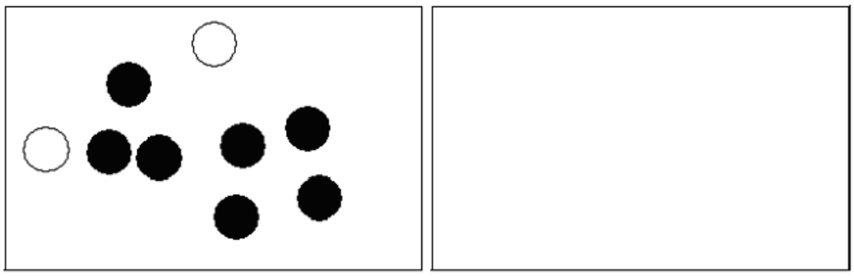
3. Two-premise syllogistic tasks.

All the experiments were conducted in Polish.

2.2.1 Picture test

As a pre-experimental research we checked semantics of "most". That is, we checked whether subjects understand this quantifier in the assumed way, namely as "more than half", i.e. according to Definition 11; analogously – for "mostnot" meaning "less than half", thus according to Definition 14.

The following experiment was planned and conducted. Different picture-models were showed to subjects, who were asked to write down all these sentences with "most" and/or Aristotelian quantifiers which they thought were true about those pictures. Universes of a chosen kind of simple objects (i.e. dots) were used. Our models differed with respect to cardinality of the universe (there were both even and odd universes) and with respect to the cardinality of objects of a given color (i.e. black and white dots). Subjects were asked to generate all sentences of a given form that were true about a given picture. First a tutorial example was presented to subjects.



Looking at the above picture you can say that:

(a) Most dots are black.

(b) Some dots are black.

(c) Some dots are not black.

Since these sentences are true about the picture:

However you cannot say that:

(d) All dots are black.

(e) Most dots are not black.

(f) No dots are black.

Since these sentences are not true about the picture.

Subsequently subjects were asked to complete further pictures:

Look carefully at the below pictures. There are 20 of them. In the box next to the picture write down all the sentences (of form a,b,c,d,e or f) you think are true about the picture. The sentences should begin with words: "most", "all", "some" *or "no" and can be both positive and negative. Please note that there can be more than one true sentence. Try to write down all the sentences you think are true. If you think that no sentence is true, draw a line.*

The following pictures were shown to subjects:

1. Even universe - 10 (5, 4, 6, 1 black dots)
2. Even universe - 8 (5, 4, 3, 6 black dots)
3. Odd universe - 9 (5, 4, 0, 2, 8 black dots)
4. Odd universe - 11 (9, 5, 11, 6, 4 black dots)

The above exhausts in our opinion all interesting cases, since for both odd and even universes there are so-called border cases: *half, one more than half, one less than half, all, no* and some additional "non-border cases" also. We used various cardinalities and made each picture slightly different with respect to the way the elements were scattered. The purpose was that subjects had no possibility to learn during the experiment *how many* "most" is in a given model and validity-checking for the quantifier "most" was not replaced with validity-checking for some quantifier $\exists^{<n}$ or $\exists^{>n}$ elements. The tasks were ordered randomly, so the cardinalities changed from picture to picture. As we had as many as twenty examples in general, it should be sufficient for statistical significance.[4]

Finally, we can group these tasks in the following disjoint categories:

Category 1 The tasks with *more than half* but *not all* dots black, so allowing "most"- (and what follows also "some"- and "somenot"-sentences) but not allowing "all"-sentences (7 "most"-models).

Category 2 The tasks with *less than half* but still *some* dots black, so allowing "mostnot"- (as well as "some"- and "somenot"-sentences) but not allowing "no"-sentences (7 "mostnot"-models).

Category 3 The tasks with *all* dots black, so allowing "all"-, "most"- and "some"-sentences (2 "all"-models).

Category 4 The tasks with only white dots, so *no* black dots, hence allowing "no"-, "mostnot"- and "somenot"-sentences (2 "no"-models).

Category 5 The tasks with *equal* number of black and white dots, so allowing *only* particular conclusions (2 tasks).

This test and the text materials from the remaining two parts of the experiment are presented in their English translation in Appendix B.

[4]Of course this is a minimal version. We could not use very many tasks since our subjects were asked to *write down* the whole sentences which requires much more time than just marking *yes/no* replies.

2.2.2 Choice of terms

Connection between so-called existential import and ontological commitment is checked by comparing subjects' readiness to generate particular conclusions on the basis of universal premises in the case of empty or non-empty terms. Let us recall, that by *empty terms* we understand terms without reference in the real world. (See also section 1.4.)

Bucholc (2004) divided terms in four groups:

1. Obviously empty ("centaur", "dwarf", "elf")
2. Obviously non-empty ("lawyer", "philosopher", "cat")
3. Empty but not in the obvious way ("the author of this paper's brothers")
4. Non-empty but not in the obvious way ("wyracs" – does an average subject know what kind of animals are "wyracs"?)

We are of the opinion that this division is artificial, since the real emptiness of a term is not important at all. It is only the epistemological aspect that matters here – namely what a subject knows (or thinks) about the ontological status of these alleged entities. Thus in this sense the difference between the third and fourth group of terms is irrelevant. These are simply terms emptiness of which is difficult to evaluate, and that is why they can be attributed to neither the first nor the second group. Furthermore it may turn out that, from the epistemological point of view, even the distinction between the first and the second group is not so obvious and natural as it may seem at the beginning.[5] Terms like "elves" and "gnomes" appear to be treated as denoting some beings appearing in literature, myths, fairy-tales and existing in a kind of imaginary sense. But when we follow this path it may turn out that queens of Poland and dinosaurs exist just in a different (past) point of time. Talking frogs and flying cars are similar to elves and gnomes – they can be imagined or appear in literature. "Present kings/queens of France/Poland" might seem to be obviously empty terms, but of course this is so obvious only for those smart people who know that these countries are now republics. The moral to be drawn from this short philosophical story is that it is very hard to determine whether terms we use in our examples are really empty for our subjects or not. From this point of view our experiment checks not only the existential import of quantifiers but also a subject's inclination to reification, and this is something that cannot be avoided. What we tried to do in our experiment

[5]To illustrate what we mean we could tell an anecdote about a situation from a logic class for first-year students, who when asked what they think about emptiness of such expressions as "an orphan Mary" or "a gnome" (there is a famous polish story for children "The orphan Mary and the seven gnomes") agreed that there are at least seven gnomes in the world, although, without understandable reason, they claimed that for sure there is no such person as "an orphan Mary", since this is a novel-character only.

was to choose terms that are empty in quite (in our opinion) an apparent way. These terms form the following three groups:

1. Terms that do not appear in language and sound strange enough so that it seems indubitable that they do not denote anything existing in the world ("pteroklaki", "mermogliny").

2. Contradictory terms - namely terms that contain a kind of factual ("flying ostrich", "singing cat") or logical ("square circle") contradiction.

3. Terms denoting fictitious beings like "Martians" which are often used in language almost as synonyms of something not existing. (We try however to avoid terms denoting beings which are very popular in literature or myths, since they seem to involve too strong reification.)

Here the first group is crucial as it presumably involves the weakest reification. In the experiment with direct inferences we used *only* terms from the first group (plus one example with "present queen of Poland"). In the two-premise experiment at least one term in each syllogism was taken from the first group.

We used the following new-invented terms in the empty-terms group: *szaruchy, mermogliny, mroczniaki, mgłowce, wietrzydła, grombliny, pupsziątka, trakloki, Zarkotrytki, leprokraki, klumpiaki.*

As middle terms in the two-premise tasks the following empty terms were used (however terms denoting fictitious beings as "gnome" or "goblin" were never used as the only empty terms in syllogisms): *flying quadrupeds, singing cats, gnomes, creatures living on Mars, sandmen, flying elephants, talking frogs, living queens of Poland, wizards, square circles, witches, beings living on Venus planet, goblins, swampmen, creatures from Never-planet.*

2.2.3 Direct inferences

For the direct inferences part, two sets of 12 sentences (with empty and non-empty terms) were prepared. Thus there were two examples for each of 6 categorial sentences: *All A's are B, No A's are B, Some A's are B, Some A's are not B, Most A's are B, Most A's are not B*. The examples were randomly ordered. Subjects were asked to generate from given sentences all the possible (if any) inferences of a given form, concerning a given relation, e.q. "relation between elephants and green colour". No tutorial example was given, so that subjects were not "taught" how they should understand quantifiers. The instruction was as follows:

> *Below you can read 12 sentences. You can infer some other sentences from them. Under each sentence write down all the sentences that in your opinion follow from the given one. Your sentences should begin with words "all", "most", "no", "some" and can be both positive or negative. There can be more than one such*

> *inference for each given sentence. If you think that you cannot infer anything, draw a line.*

Additionally, to hinder subjects from basing on their knowledge about world, it was explicitly told in the control groups that they should not worry if some sentences seem to them absurd or false in the light of they know about the world:

Such a remark was not given in the empty terms group. We were interested in subjects' inferences in the case where premises were about non-existing objects, so that we could estimate on the basis of these inferences what kind of interpretation subjects give to such sentences. That is why we did not want to suggest or impose the interpretation that such premises are false, or may be understood as false.

2.2.4 "Most" working in complex reasoning: two-premise inferences.

The choice of moods

In the last part of the the experiment, two-premise inferences were tested, again in two groups: with empty and non-empty terms. For this purpose 20 pairs of premises out of 21 + 10 (namely of all the possible valid syllogisms (21) with "most" plus quasi-inferences dependent on the implicature of the "most"-premise (10) – see chapter 1.5) were chosen. We took almost *all the combinations of premises* which vary in any essential sense, omitting only the ones that are in a sense equivalent. For instance the pair $All(B, A)$ and $Most(C, B)$ (I Fig.) is equivalent in a sense of possible inferences to the pair $Most(A, B)$ and $All(B, C)$ (IV Fig.), and hence we used only the first one in our experiment.

The chosen pairs of premises can be divided into five disjoint categories:

Category 1 Premises allowing "most/mostnot"-conclusions (and hence also particular ones), but not allowing any universal conclusions (1,2,7,8).

Category 2 Premises allowing *only* (classical)[6] particular conclusions (16,17,13,20).

Category 3 Premises allowing universal conclusions (and as well "most/mostnot" and particular ones with the assumption of non-empty domains taken into account) $(5, 6, 11)$.

Category 4 Premises allowing only implicature-based particular conclusions $(3, 4, 9, 10, 15, 18, 19)$.

Category 5 Premises allowing only both together: a classical particular conclusion and a conclusion dependent on the implicature, e.g. from $Most(A, B)$

[6]not implicature-based

and $\mathit{All}(B, C)$ we can conclude: $\mathit{Some}(C, A)$ and $\mathit{Somenot}^*(C, A)$ (the latter is dependent on the implicature of the premise with "most"), and from $\mathit{Mostnot}(B, A)$ and $\mathit{All}(B, C)$ we can conclude $\mathit{Somenot}(C, A)$ and also $\mathit{Some}^*(C, A)$ $(12, 14)$.

The following moods were chosen for the experiment:[7]

1. Fig. I

$$\begin{array}{c}(1)\\ \mathit{All}(B,A)\\ \mathit{Most}(C,B)\\ \hline \mathit{Most}(C,A)\\ \mathit{Some}(C,A)\end{array}\quad \begin{array}{c}(2)\\ \mathit{No}(B,A)\\ \mathit{Most}(C,B)\\ \hline \mathit{Mostnot}(C,A)\\ \mathit{Somenot}(C,A)\end{array}\quad \begin{array}{c}(3)\\ \mathit{All}(B,A)\\ \mathit{Mostnot}(C,B)\\ \hline \mathit{Some}^*(C,A)\end{array}\quad \begin{array}{c}(4)\\ \mathit{No}(B,A)\\ \mathit{Mostnot}(C,B)\\ \hline \mathit{Somenot}^*(C,A)\end{array}$$

$$\begin{array}{c}(5)\\ \mathit{All}(B,A)\\ \mathit{All}(C,B)\\ \hline \mathit{All}(C,A)\\ \mathbf{Most}(\mathbf{C},\mathbf{A})\\ \mathbf{Some}(\mathbf{C},\mathbf{A})\end{array}\quad \begin{array}{c}(6)\\ \mathit{No}(B,A)\\ \mathit{All}(C,B)\\ \hline \mathit{No}(C,A)\\ \mathbf{Mostnot}(\mathbf{C},\mathbf{A})\\ \mathbf{Somenot}(\mathbf{C},\mathbf{A})\\ \mathbf{Mostnot}(\mathbf{A},\mathbf{C})\\ \mathbf{Somenot}(\mathbf{A},\mathbf{C})\end{array}$$

2. Fig. II

$$\begin{array}{c}(7)\\ \mathit{No}(A,B)\\ \mathit{Most}(C,B)\\ \hline \mathit{Mostnot}(C,A)\\ \mathit{Somenot}(C,A)\end{array}\quad \begin{array}{c}(8)\\ \mathit{All}(A,B)\\ \mathit{Mostnot}(C,B)\\ \hline \mathit{Mostnot}(C,A)\\ \mathit{Somenot}(C,A)\end{array}\quad \begin{array}{c}(9)\\ \mathit{All}(A,B)\\ \mathit{Most}(C,B)\\ \hline \mathit{Somenot}^*(C,A)\end{array}\quad \begin{array}{c}(10)\\ \mathit{No}(A,B)\\ \mathit{Mostnot}(C,B)\\ \hline \mathit{Somenot}^*(C,A)\end{array}\quad \begin{array}{c}(11)\\ \mathit{All}(A,B)\\ \mathit{No}(C,B)\\ \hline \mathit{No}(C,A)\\ \mathbf{Mostnot}(\mathbf{C},\mathbf{A})\\ \mathbf{Somenot}(\mathbf{C},\mathbf{A})\\ \mathbf{Mostnot}(\mathbf{A},\mathbf{C})\\ \mathbf{Somenot}(\mathbf{A},\mathbf{C})\end{array}$$

3. Fig. III

$$\begin{array}{c}(12)\\ \mathit{Most}(B,A)\\ \mathit{All}(B,C)\\ \hline \mathit{Some}(C,A)\\ \mathit{Somenot}^*(C,A)\end{array}\quad \begin{array}{c}(13)\\ \mathit{No}(B,A)\\ \mathit{Most}(B,C)\\ \hline \mathit{Somenot}(C,A)\end{array}\quad \begin{array}{c}(14)\\ \mathit{All}(B,A)\\ \mathit{Mostnot}(B,C)\\ \hline \mathit{Some}^*(C,A)\\ \mathit{Somenot}(C,A)\end{array}\quad \begin{array}{c}(15)\\ \mathit{No}(B,A)\\ \mathit{Mostnot}(B,C)\\ \hline \mathit{Somenot}^*(C,A)\end{array}$$

$$\begin{array}{c}(16)\\ \mathit{Most}(B,A)\\ \mathit{Most}(B,C)\\ \hline \mathit{Some}(C,A)\end{array}\quad \begin{array}{c}(17)\\ \mathit{Mostnot}(B,A)\\ \mathit{Most}(B,C)\\ \hline \mathit{Somenot}(C,A)\end{array}$$

4. Fig. IV

[7] We write in bold those inferences which require an additional assumptions that domains are non-empty.

Category	Fig. I	Fig. II	Fig. III	Fig. IV
I	1,2	7,8	-	20
II	-	-	16,17,13	-
III	5,6	11	-	-
IV	3,4	9,10	15	18,19
V	-	-	12,14	-

Table 2.1: Classification of the moods

$$\begin{array}{c}(18)\\ No(A,B)\\ Mostnot(B,C)\\ \hline Somenot^*(C,A)\end{array}\qquad\begin{array}{c}(19)\\ Mostnot(A,B)\\ All(B,C)\\ \hline Some^*(C,A)\end{array}\qquad\begin{array}{c}(20)\\ No(A,B)\\ Most(B,C)\\ \hline Somenot(C,A)\end{array}$$

Table 2.1 shows which of the above moods belong to each of the categories.

The experimental set-up

The experimental set-up consisted of a tutorial and 20 sets of premises randomly ordered. (Thus each syllogism was used only once in each of the two groups.) In each case we asked a subject what (if anything) she could tell on the basis of given sentences about mutual relations between given entities. Because of the difficulty of the two-premise inferences, subjects were instructed how they should understand a task. This was done by presenting a simple tutorial example, but none of the really controversial syllogisms, so subjects were also not "taught" about how to infer implicatures or so-called existential import.

Tutorial:

Below you can read 20 *examples of different pairs of sentences. From these pairs of sentences you can infer another sentences. For instance if you know that:*

All dwarves have beards.
All gnomes are dwarves.

You can tell that:

All gnomes have beards.

Instruction:

The instruction was similar to the one in the direct inferences part, however some elements were more stressed – e.g. that it is possible of infer more than one sentence for each pair of premises. Subjects in the control group were told to reason only on the basis of given premises:

> *Some of the sentences may seem to you false or absurd in the light of your knowledge about the world, e.g. "All dentists like mandarins". Do not worry about it. Try to reason only on the basis of the given sentences and not on the basis of your knowledge.*

Each example in the test was different with respect to the content – each was about different entities. The question specified in each case what kind of relation should be considered by the subject (e.g. *What (if anything) can you tell on the basis of these sentences about mutual relations between gnomes and having beards?*)

2.3 Subjects

Table 2.2 summarizes the number of subjects in groups, their age and field of study. We tested library science and philosophy students. Although a lack of logical background was recommended, the complexity of the tasks (especially in the case of the two-premise part) required a minimal understanding of what is asked, e.g. what it means to infer a conclusion. In this respect philosophy students, due to having basic logical skills, seem to be slightly better than average. In all parts of the experiment, subjects were given 40 minutes to complete tests with 20 tasks (in practise usually c.a. 30 minutes was enough). They were tested in groups, however without the possibility of communicating with each other. They were also told that the experiment is no exam nor IQ-test and that its purpose is to check people's intuitions.

GROUP	**Picture-test**	**Direct Inferences**		**Two-premises part**	
		EG	**CG**	**EG2**	**CG2**
number of subjects	24(22*)	25	24	24	26
age average	21	21.9	2.,3	21	20.2
age range	20-22	20-24	20-22	18-25	19-25
field of study (year)	library science (III)	philosophy, library science (III)	library science (III)	philosophy (I)	philosophy (III)

Table 2.2: Subjects

* – 2 tests were excluded because of subjects' misunderstanding of the task, so there were 22 valid tests.

2.4 Results: picture test

The results were quite apparent. The experiment showed clearly that subjects understand "most" as meaning "more than half" ("mostnot" – meaning "less than half"). Moreover, the quantifier "most" is not used in cases when sentences with "all" are true, (analogously "mostnot" when "no"-sentences are true), which we connect with scalar implicature of these quantifiers. Additionally we obtained some interesting results concerning the quantifier "some". Below we present the detailed results grouped by categories of tasks (see 2.2.1).

2.4.1 Categories 1 and 2

As far as sentences with the quantifier "most" are concerned the results were as follows. For **Cat.1** 94.8% of subjects generated a sentence with "most" (variance with respect to task was $v = .8$), whereas for **Cat.2** 92.8% of subjects generated a sentence with "mostnot" (variance: $v = 1.95$). The results for "some"- and "somenot"-sentences (with variances) for both categories are summarized in Table 2.3.

Basing on the results, we conclude that subjects clearly understand the quantifier "most" (resp. "mostnot") in the assumed way, so that it means "more than half" (at least in small domains). Such a conclusion is allowed since:

1. Our tasks contained all the possible border-cases and almost 100% of subjects generated correct sentences with "most" (resp. "mostnot") for all the relevant models.

2. The variances for both "most"- and "mostnot"-sentences were rather low, so subjects' correctness did not depend on the particular task.

3. The "most/mostnot"-sentences almost did not occur at all in the case of models in which they would be false. "Most" never occurred for an improper model and "mostnot" was generated falsely only 9 times for all $242 = 11(tasks) * 22(subjects)$ cases, which gives 3.7%. [8]

It is worth noting that in the case of particular conclusions, a significantly higher percentage of subjects were inclined to generate the "implicature-based" particular sentence (namely "somenot" for "most"-models and "some" for "mostnot"-models) out of the two *true* particular sentences. Thus for example for a model with 9 black dots in the universe of 11 dots more subjects responded "Some dots are not black" than "Some dots are black". (See table 2.3.)

[8] 3 out of these 9 cases occurred for models of category 5 - namely with equal cardinality of white and black dots. There are two possible reasons for these fallacies: (a) counting mistakes and (b) doubts concerning the proper semantics of "mostnot" – whether it means "less than half" or "not most".

category	*some*	*somenot*
1. – "most"-models	59% $v = 1$	88% $v = .95$
2. – "mostnot"models	91.5% $v = 8.4$	62.3% $v = 7.9$

Table 2.3: Difference between "some" and "somenot" senetences generated for "most"- and "mostnot"-models in the picture test.

The significance of this result was proved both for "some"- and "somenot"-sentences.[9] We checked whether there is a significant difference in the mean number of sentences with "some" (dependent variable *some*), resp. "somenot" (dependent variable *sn*), generated in the whole group of 24 subjects for the two categories of tasks: category 1 and category 2.

We expected the number of "some"-sentences (*some*) to be significantly larger in Cat.1. than in Cat.2. and the number of "somenot"-sentences (*sn*) significantly smaller in Cat.1. than in Cat. 2.

An exploratory analysis was first conducted to check if the data are parametric. Both Kolmogorov-Smirnov and Shapiro-Wilk tests showed a significant deviation from normality for a dependent variable *some* (in both independent groups). Although the normality tests were not significant for *sn* – *p*-values were higher than the critical $\alpha = 0.05$, they were also ≤ 0.2 which with such small sample sizes ($N = 14$; 7 in each independent group) cannot exclude the deviation from normality. Levene's tests for homogeneity of variances in the two compared groups were not significant for either dependent variables *some* and *sn*, which suggests that the assumption of equal variance was satisfied, however the *p*-values were close to significance and accounted to $p = .07$ for *some* and $p = .197$ for *sn*.

Basing on the exploratory analysis, we decided to use non-parametric tests for checking the between-group difference. A Mann-Whitney test was conducted for both *some* and *sn* and proved a clearly significant between-group difference for both dependent variables.

For *some*: The number of "some"-conclusions generated for "most"-models (Cat.1) ($Mdn = 13$)[10] was significantly smaller than the number of "some"-conclusions generated for "mostnot"-models (Cat.2) ($Mdn = 21$): $U = 0$; $z = -3.057$, $p < .001$,[11] the effect size: $r = -.817$.

For *sn*: The number of "somenot"-conclusions generated for "most"-models (Cat.1) ($Mdn = 20$) was significantly larger than the number of "somenot"-conclusions generated for "mostnot"-models (Cat.2): ($Mdn = 14$) $U = 1$; $z =$

[9] SPSS 15.0 was used for all the analyses.

[10] $Mdn = median$

[11] In both cases we took the *exact* (small sample sizes) one-tailed (directorial hypothesis) significance into account.

-3.155, $p < .0001$, the effect size: $r = -.843$.[12]

Conclusion 1. *The percentage of people who tend to generate sentences with "some" is significantly lower if the number of objects to which these sentences refer is smaller than half of the whole domain than if the number of these objects is larger than half. The converse is true for sentences with "somenot". Since models are given, and both sentences are true in both categories, to explain the above result we have to consider scalar properties of the quantifiers "some/somenot". (More on this issue in the part 2.4.3.)*

2.4.2 Categories 3 and 4

For those categories, sentences with "most/mostnot" or "some/somenot" were generated very rarely. Only 2 out of 22 subjects generated a conclusion with "most" for a model with all dots black, whereas the universal conclusions were generated by a high percentage of subjects: 93.18% for "all"-models (Cat.3.) and 90.9% for "no"-models (Cat.4.) (variances for this tasks were: $v = 0$ for "no"-sentences and $v = 0.5$ for "all"-sentences).

Apparently people are not inclined to use the quantifiers "some/somenot" or "most/mostnot" in the case of models for which universal statements are true. They either understand these quantifiers as false in such models (which would imply that the implicature "not all" is rooted even in the semantics of these quantifiers), or they just do not regard uttering sentences with these quantifiers as normal language behavior because of pragmatic reasons. If it is possible to give a stronger sentence, then it is required by the conversational rules, and then also the weaker expression may seem unnecessary (especially given that multiple replies were rather rare and occurred only for "most"- and "mostnot"-models). In any case we have a strong premise to conclude that people understand these quantifiers as *at least* implicating "not all/not no".

2.4.3 Particular conclusions for joint categories: 1, 2 & 5

The results for particular sentences for *joint categories of tasks* in which we could expect this kind of response (so excluding "all"- and "no"-models in which such responses did not occur at all) are characterized by quite high variance: $v = 16.5$ for "some"-sentences and $v = 11.9$ for "somenot"-sentences. We observed that the number of subjects generating the particular sentence "Some dots are black" decreased with the number of black dots on the picture. (The opposite effect was observed for "somenot"-conclusions). We already proved a significant difference in the number of generated "some/somenot"-sentences between tasks of Cat. 1. and Cat. 2. We further suspected a general correlation between

[12] To compare, parametric tests had similar results. For *some*: $t(12) = 6.139$; $F(1, 12) = 37.688$; $p < .0001$. For *sn*: $t(12) = -5.207$; $F(1, 12) = 27.113$; $p < .0001$.

		some	*sn*
bdots	Pearson Correlation	-.493	.800
	Sig. (1-tailed)	.026	.000
	N	16	16
proc	Pearson Correlation	-.564	.841
	Sig. (1-tailed)	.012	.000
	N	16	16

Table 2.4: Correlation between the number of "some/somenot"-sentences and number/precentage of black dots in the model

the number of black dots and the number of "some"- (negative correlation) or "somenot"-sentences (positive correlation). The other possible correlation could be between the number of these conclusions and the *percentage of black dots in the universe* (which we will call a "model-dependent cardinality"). As we have already mentioned, both border-cases were excluded:

1. no black dots in the model,

2. all dots black.

In the first case a sentence "Some dots are black" is false and a sentence "Some dots are not black" does not appear because of the implicature effect ("somenot" as "not no"). On the other hand, in the second case a sentence "Some dots are black" does not appear because of the implicature ("some" as "not all") and "Some dots are not black" is just false.

We checked whether the following correlations are significant:

- (negative) correlation between the variable *some* and the number of black dots (*bdot*) and/or the percentage of black dots (*proc*) in the universe.

- (positive) correlation between the variable (*sn*) and the number of black dots (*bdot*) and/or the percentage of black dots (*proc*) in the universe.

According to our predictions, we proved a strong correlation (on the level of $\alpha = .05$) in all four cases. (See Tables 2.4 and 2.5.) The correlation coefficients were higher and the significance stronger ($p < .01$) for non-parametric tests. (We need this statistic at least for the variable *some* which shows a deviation from a normal frequency distribution.)

Further on we were interested in the nature of these dependencies. We carried out four different liner regression analyses to determine whether there is a linear relationship in all the above-mentioned cases. [13]

[13]We are aware that our data are not appropriate for linear regression, especially *some* which is clearly non-parametric. Also the sample size is too small and thus the below statistics

		some	*sn*
	Spearman's rho	-.598	.820
bdots	Sig. (1-tailed)	.007	.000
	N	16	16
	Spearman's rho	-.736	.849
proc	Sig. (1-tailed)	.001	.000
	N	16	16

Table 2.5: Correlation between the number of "some/somenot"-sentences and number/precentage of black dots in the model – non-parametric

We regressed both *some* and *sn* on *bdot*, or alternatively on *proc*, which served as independent variables. For the linear regression of *some* on *bdot*, ANOVA was not significant ($p = .053$) and the assumptions of the analysis were violated. The three remaining regressions: *some* on *proc*, *sn* on *bdot* and *sn* on *proc* were significant ($p < .05$), however they were difficult to interpret because the assumptions were again violated. For more detailed reports of these analyses see Appendix A.

Because of violation of the important assumptions of linear regression we cannot conclude that there is a linear dependence in the above cases. This situation can be caused by too small sample sizes, but can be also a signal that the considered relationship is indeed non-linear. Curve estimation analysis was conducted to find out whether some other non-linear relationships are more adequate to model those dependencies. It turned out that for both *some* dependence on *proc* and on *bdot*, *cubic regression* made the best fit, whereas the dependence of *sn* on both predictors was better described by *linear* or *logarithmic* relation. (See Appendix A.)

One may also make conclusions based on analyzing graphs. Graph of the dependence of *some* on *bdot* (Figure 2.1) shows clearly that the score of the dependent variable is very low (but different from 0) for the value of $bdot = 1$ (we know also that for $bdot = 0$, which was an excluded case, $some = 0$), then it increases rapidly having the highest values for $bdot =$ 2,3 and 4, and then begins to decrease. Similarly *some* depends on *proc*. The scores of *some* are first low (for $proc = 10$), then increase having highest values in the range between 20 and 40, and then they decrease (Figure 2.2).

For dependencies of *sn* on *proc* and on *bdot* the curve estimation proved non-significant coefficients for cubic regression, whereas the linear regression was highly significant. Graphs (Figures 2.3 and 2.4) clearly show that *sn* scores increase regularly with the scores of *bdot* and similarly with the scores of *proc*. Still we have to remember that the border case of all dots black in the model was

should be treated carefully. We are however interested is the preliminary estimation of plausible dependencies.

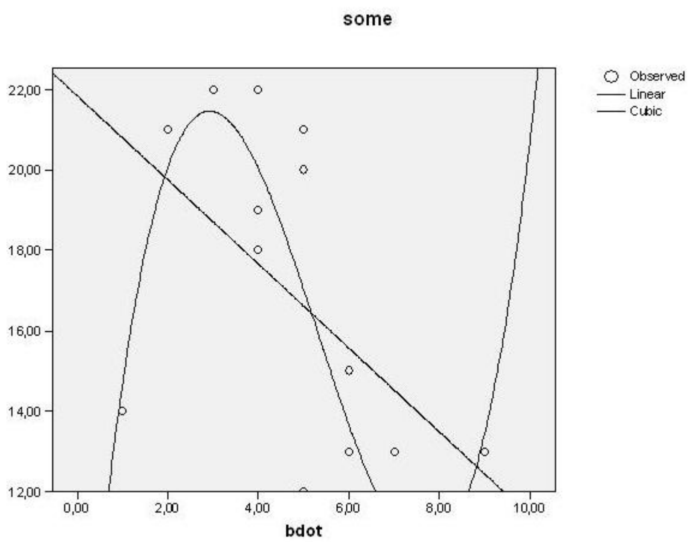


Figure 2.1:

excluded. Otherwise we would observe the sudden decline (to 0!) of the scores of *sn* at some point and probably we could obtain a significant coefficient for polynomial regression.

Since all the statistics for linear relationship were significant (on the level $p < .0001$) and as well the graphs proved its adequacy we could have agreed with such a model of dependencies of *sn* on *proc* and *sn* on *bdot* and explain the lack of satisfied assumptions by too small sample sizes. However, a similarly high significance was proved in those cases for *logarithmic* relationship. All the tests were significant on the level of $p < .0001$, and the explained variance was even higher for logarithmic models than for linear ones. It seems plausible that in reality the considered dependencies are not linear but the values of the dependent variable *sn* grow with the values of *proc/bdot* variables slower than linearly.

Summarizing, a subject's inclination to generate "some"-sentences as true of given objects (measured by a number of subjects who generated such a sentence in the whole group) equals 0 for models with 0 elements of a referred quality, is low for models with only 1 such an object, then increases rapidly having highest values for 2−3 such objects and then decreases (even to 0 for models in which all elements have a referred quality). It works similarly if we consider not a cardinality of referred objects but a percentage of them in a universe. On the other hand, a subject's inclination to generate "somenot"-sentences grows monotonically with the number of referred objects (provided that the model does not contain only this kind of elements) or alternatively with the percentage of objects of a referred quality in a universe (but of course would decrease to 0 for 100%). This growth seems to be, however, slower than linear and thus a logarithmic regression makes

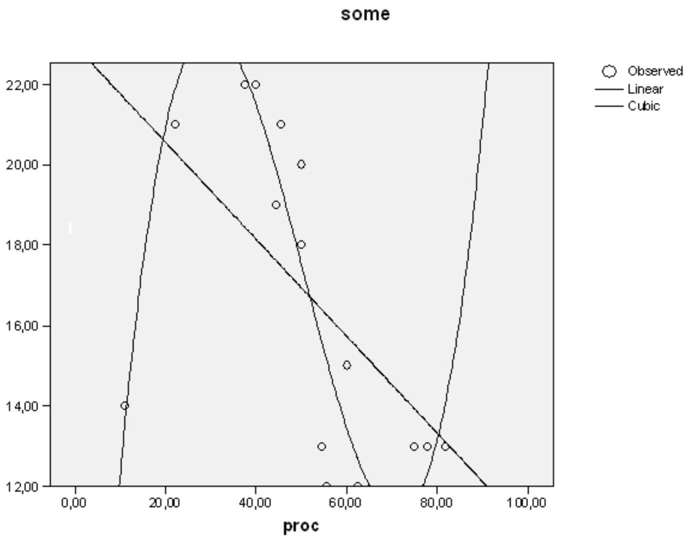


Figure 2.2:

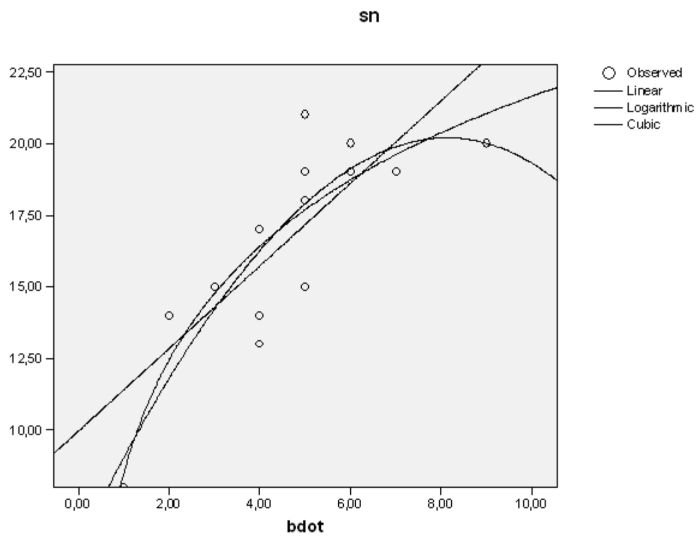


Figure 2.3:

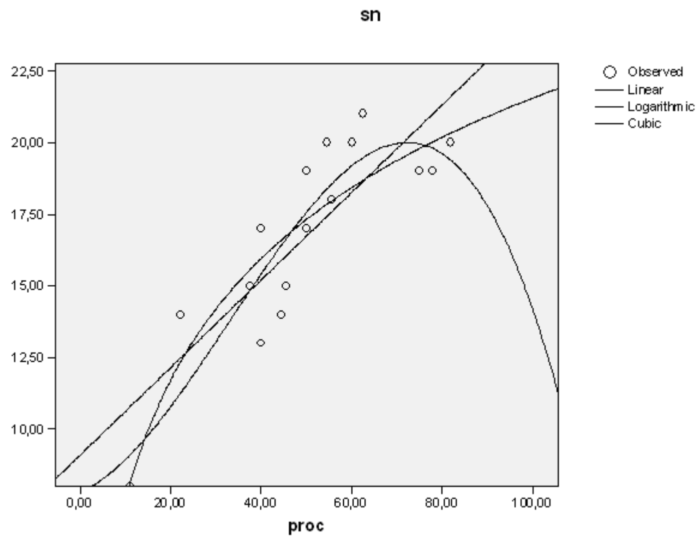


Figure 2.4:

probably a better fit.

The above analysis gives us a strong base for presumptions concerning a kind of scalar implicature of the quantifier “some”. According to the above results we can conclude that “some” means “not too many” (or “not too many in a given domain”) and thus it is less willingly used when referring to bigger samples of objects (or bigger subsets of fixed domains). This “cardinality dependence” of the quantifier “some” is systematic.

Since our experiment was not oriented towards investigation of this kind of relationships, we suspect that one could get rid of some of the deficiencies of the analyses with bigger samples and a more suitable selection of models. We are of the opinion that an experiment that could confirm these results and would allow us to compare the differences between the cardinality dependence (*bdot*) and the “model” cardinality dependance (*proc*) is worth taking into consideration.

A few final words concerning the critical values of the independent variable *bdot* that showed importance in our analyses. First of all number 2. It seems that “some” means “at least two”, at least in Polish. Less than half of subjects gave a response with “some” for models with only one black dot and this number increased considerably for 2 dots. This result may be connected with the plural form of a Polish quantifier “niektóre” and the sentence (in plural) given in the tutorial. “Niektóre” does not occur in Polish in singular at all.[14] The other critical value is number 4. The curves for *some* had extrema close to 4 and displayed a decline beginning from 5. A noteworthy fact may be that this very number is

[14]However we got 14 such responses for a model with one black dot, which is still a lot.

considered in cognitive science as an upper bound of human capacity of *subitizing*, thus immediate, accurate, and confident judgments of number in case of small numbers of items. It has been observed that while people are able to estimate any number of items immediately if this number is smaller than approximately 4, there is a significant growth of reaction time with each added item over 4. Thus, there appear to be real differences in the ways in which a small number of elements is processed by the visual system (i.e., approximately < 4 items), compared with larger numbers of elements (i.e., approximately > 4 items). While numbers smaller than 4 can be recognized "directly", larger numbers require probably some kind of enumeration process. (See e.g. (Taves, 1941), (Kaufman et al., 1949), (Trick & Pylyshyn, 1994).) The question is, whether we can connect our results with the subitizing phenomenon. It is possible that "some" may refer to very small numbers in the range of subitizing, but only in the case of small domains. We suppose that in bigger domains it is more dependent on how big a part of the given domain the referred number is. Thus, the semantics of "some" may be in this sense domain-dependent.

2.5 Results – direct inference

2.5.1 Working hypotheses

We specify the working hypothesis for this part.

Hypothesis 1. *The scalar implicature of "most" and "mostnot" will interfere in the process of reasoning.*

(A) "Some"-sentences will occur as conclusions from "mostnot"-premises and "somenot"-sentences as conclusions from "most"-premises.

(B) "Most" and "mostnot"-sentences will not occur significantly as conclusions of "all/no".

The point B requires some explanation. Even if due to the implicational scale "all" implies "most" and further "most" implies "some", a subject may not consider it justifiable enough to infer a weaker statement when a stronger one is already given. They may find uttering such a statement weird language behavior.

Hypothesis 2. *The quantifier "most" will display stronger so-called existential import than "all" (resp. "mostnot" than "no"). It means that we expect particular sentences to be inferred more frequently from "most/mostnot"-premises than from "all/no"-premises.*

Hypothesis 3. *Scalar implicatures will occur as inferences for premises with "some" (resp. "somenot"), namely conclusions with "somenot" and/or "mostnot" will occur for premises with "some", and conclusions with "some" and/or "most" for premises with "somenot".*

Hypothesis 4. *Basing on the general hypothesis IV, we expect expect significant between-group effects – lesser tendency for generating particular sentences or sentences with "most" as conclusions from universal sentences in the EG than in the CG.*

2.5.2 Results

1. Implicatures

   Hypotheses 1 and 3 can be sustained. We observed that for both "most/mostnot" and "some/somenot"-premises scalar implicatures were preferred by quite a high percentage of subjects. For "most/mostnot"-premises, scalar implicatures occurred also even more often than so-called existential import. There was no between-group effect.

   As far as scalar implicatures for "most" and "mostnot" are concerned (hypothesis 1) the results were as follows (Table 2.6): 76% in the EG and 70.8% of subjects in CG inferred "somenot"-conclusions from "most"-premises; for "mostnot"-premises those numbers for "some"-conclusions accounted to: 72% (EG) and 68.7% (CG).

   This result was checked with χ^2 test to be highly significant for both types of premises: with "most" and "mostnot" quantifiers, in both groups.

   For both types of premises, the χ^2 tests showed a significant difference in the number of subjects who generated implicatures for both examples (each premise was repeated twice), one of them or not at all.

   EG: For "most" $\chi^2(2) = 17.36$, $p < .0001$; for "mostnot" $\chi^2(2) = 14.48$, $p = .001$.

   CG: For "most" $\chi^2(2) = 9.25$, $p = .012$; for "mostnot" $\chi^2(2) = 6.75$, $p = .041$.

   Additionally, the histograms showed a clear negative skew, which means that most subjects scored 2, so generated implicature coherently for each of the questions of a given form. We conclude that according to our predictions, implicatures were generated for "most" and "mostnot" premises on a significant level.

   To compare, the existential import for both "most" and "mostnot" had much lesser effect: 58% (EG) and 47.9% (CG) for "most", and 46% (EG) and 27% (CG) for "mostnot". (See table 2.9.)

   Additionally, although "most/mostnot"-sentences occurred as conclusions from "all/no"-premises, they were the least frequently inferred conclusions for universal premises. "Most" was inferred from "all" by 44% of subjects in EG and only 12.5% in CG, whereas "mostnot" from "no" by 24% in EG and only 6.25% in CG (see table 2.8). This result contradicts hypothesis

Quantifier in a premise	Conclusion	EG(%)	CG(%)
Most	Somenot	76	70.8
	Some	58	47.9
	Most	16	4.1
Mostnot	Some	72	68.75
	Somenot	46	62
	Mostnot	18	8.3

Table 2.6: Scalar and existential implicature of "most" and "mostnot"

Quantifier in a premise	Conclusion	EG (%)	CG (%)
Some	Somenot	58	35.4
	Mostnot	34	27
	Some	28	31.25
Somenot	Some	52	45.8
	Most	34	22.9
	Somenot	34	14.6

Table 2.7: Scalar implicature of "some" and "somenot"

4, as conclusions with "most" were more frequently inferred from universal premises in the empty-terms group than in the control group.

Results for the similar kind of scalar implicature ("not all/not no") (hypothesis 3) of "some/somenot"-premises were less apparent (see table 2.7). Although implicatures were again the most willingly inferred sentences, "somenot"-conclusions were inferred from premises with "some" by only 35.4% of subjects in the control group and 58% in the empty-terms group, while "some"-conclusions were inferred from "somenot"-premises by 52% of subjects in the EG and by 45.8% in the CG. This low result is connected with the low rate of *any* conclusions generated for this category of task, which was the one most frequently left without any response.

"Mostnot"-sentences occurred even less as implicatures of premises with "some": 34% (EG) and 27% (CG), whereas sentences with "most" were inferred from "somenot"-premises by 34% (EG) and 22.9% (CG).

Finally there was no significant between-group (EG vs CG) difference (on the basis of Mann-Whitney tests).

2. Results for universal premises ("all", "no") and Hypothesis 2 .

   The most frequent inferences for both universal quantifiers turned out to be the universal ones with the use of the other universal quantifier – so-called *obverse* forms, namely "All A's are not B" (e.g. "All elephants are not green") as a conclusion from the "no"-premise: "No A's are B" ("No elephants are green"), which was inferred by 64% of subjects in the CG and by 52% in the EG, and "No A's are not-B" (e.g. "No doctor is without a hat") as a conclusion from the "all"-premise: "All A's are B" ("All doctors have hats"), which was inferred by 20% of subjects in the CG and 64% in the EG. (See Table 2.8.) The above result shows that "no" and "all" are similarly understood in Polish, as the universal quantifier. (This fact can be also connected with the double negation in the Polish language.[15])

Hypothesis 2 can be sustained. Subjects inferred so-called existential import from universal premises less frequently than from premises with "most". Conclusions with "some" were inferred from "all"-premises by 25% in CG and 52% in EG, whereas conclusions with "somenot" from "no"-premises by 8.3% (CG) and 28% (EG). Each of these numbers is smaller than the corresponding number for "most" (see also Tables 2.6 and 2.9).

To test Hypothesis 2 we checked whether there is a significant difference in the mean result for the so-called existential-import generated for the four categories of premises – with quantifiers: "most", "mostnot", "all" and "no". For each category the scores ranged between 0 and 2 (0 – lack of so-called existential import; 1 – so-called existential import generated for one of the tasks; 2 - so-called existential import generated for both tasks). Because of high deviation from normality (on the basis of Kolmogorov-Smirnov and Shapiro-Wilk tests), we conducted a non-parametric procedure. The Friedman test was conducted. The weight of existential import significantly changed depending on the condition (so the quantifier in the premise). We report the results of the test for the empty-terms group and the control group. For EG: $N = 25$, $\chi^2(3) = 10.487$, $p = .012$; for CG: $N = 24$, $\chi^2(3) = 17.257$, $p = .000$.

Subsequently, a post-hoc procedure was conducted to check which of the four conditions differ significantly. Wilcoxon tests were used for this purpose. Three comparisons were made: "most" – "mostnot", "mostnot" – "all", "all" – "no". A Bonferroni correction was applied, so all effects are reported at a level of $\alpha = .0167$.

For both groups only the comparison "no" – "all" turned out to be significant. We report exact 1-tailed significance since our hypothesis is directorial and assumes lower so-called existential import for universal and negative quantifiers.

For EG: $z = -2.585$, $p = .004$, $T = 0$, the effect size: $r = -.365$.

[15]The Polish universal negative sentence "Żadne A nie sa B" can be literally translated to English as: " No A-s are *not* B". So the preferred conclusion from "all"-premise sounds in fact as "No A's are *not* not-B" and the sentence "Wszystkie A nie sa B" (lit. "All A's are not B") has the same semantics as "Żadne A nie sa B" (lit. "No A's are *not* B").

Quantifier in a premise	Conclusion	EG (%)	CG (%)
All	No... not	64	20
	All	16	4,1
	Some	52	25
	Most	44	12.5
No	All... not	52	64
	No	24	18
	Somenot	28	8.3
	Mostnot	24	6.25

Table 2.8: Results for universal quantifiers

Quantifier in a premise	EG(%)	CG(%)
Most	58	47.9
All	52	25
Mostnot	46	27
No	28	8.3

Table 2.9: So-called existential Import

For CG: $z = -2.271$, $p = .016$, $T = 0$, the effect size: $r = -.327$.

Conclusion 2. *We conclude that subjects displayed significantly decreasing so-called existential import due to the quantifier in the premise ($Most \longmapsto Mostnot \longmapsto All \longmapsto No$) however the significant effect one obtains only with the jump to the negative universal quantifier ("no"). Thus both use of negation and universality of a quantifier turn out to have influence on this kind of inferences, namely both negation and universality of the quantifier* weaken *existential import.*

3. Between-group differences.

Hypothesis 4 cannot be sustained. So-called existential import was even stronger in the empty-terms group, which contradicts the hypothesis. (See table 2.9.) We consider this effect to be connected with the fact that our groups were not similar enough with respect to their background. In the control group we had only library science students who have very weak or even no experience in logic. The empty-terms group contained mostly 3rd and 4th year philosophy students with some background in logic and more experience in reasoning tasks in general. The lower effect of existential import in the control group is connected also with a lower percentage of subjects in this group inferring *any* of the expected conclusions and with a higher tendency to generate mistakes or just sentences of a strange (difficult to classify) form.

Another reason for the above might be the context-dependence of reasoning with non-empty terms. The following responses appeared for example in tests.

Premise 1: *Most philosophy students wear berets.*

Conclusion: *Some philosophy students follow their lecturers' example and try to imitate them.*

Premise 2: *No politician has a bike.*

Conclusions 1: *Politicians lead a very unhealthy life. Politicians steal so they can afford cars.*

There were also many "giving-cause responses", i.e. for premise 2:

Conclusion 2: *All politicians got rid of bikes for the sake of political propaganda*

or "giving-effect responses", e.g.:

Premise 3: *All lawyers like cheese.*

Conclusions: *Some lawyers would demand cheese for dinner. All lawyers would be satisfied with cheese for dinner.*

The above-described phenomenon can be an effect of not understanding the task (which explicitly says what sort of conclusion we expect) by subjects with weaker experience in reasoning tasks or of memory-retrieval (a subject appeals to her knowledge about the world).

4. "Equivalence" inferences:

   Occasionally, we observed also that subjects inferred sentences that were logically equivalent to the given premises. Such responses occurred, however, rarely and even if they occurred subjects tried to use different words. Thus, subjects rather reformulated the premise than repeated it directly. For example instead of "Some A's are B" they would give "Some B's are A" or "Not all A's are not B".

2.6 Results – two-premise inferences

2.6.1 Hypotheses

This part should be treated more as a pilot research the aim of which was to check how people cope with different syllogistic moods with "most". For this purpose we selected all possible syllogisms which differ in any essential way. We expected the following dependencies:

1. The influence of negation in premises or/and in a conclusion on the difficulty of the task, so on the number of "correct" inferences. By a *"correct"* inference we understand here an inference that is correct in the classical sense, or it is based on implicature, or on so-called existential import.

2. Mutual dependencies between all the possible conclusions of a given set of premises:

   (a) Suppression of a particular conclusion in the presence of a universal or "most/mostnot"-conclusion (scalar implicature of "some").

   (b) Suppression of a "most/mostnot"-conclusion in the presence of a universal conclusion (scalar implicature of "most").

   (c) Suppression of an implicature in the presence of a classical conclusion.

The well-known effect of difficulty differences between syllogistic figures should also be taken into account in the analysis.

2.6.2 Results

Syllogisms of category III (universal conclusions)

In the case of these syllogisms (for the classification of categories see section 2.2.4) universal conclusions were inferred by almost 100% of subjects, whereas the other possible (with quantifiers "some" or "most") conclusions were completely ignored. This result shows no significant between-group (EG2 vs CG2) difference.

Control group (CG2): Subjects generated almost only universal conclusions (one "mostnot"-conclusion for 6 syl.[16]) Apart from this the particular and "most"-conclusions were not inferred at all. Only one mistake occurred - in case of 5 syl. one subject, except for the correct inference (All(C,A)), gave also an incorrect one - namely a *converse* form of the "all"-sentence (All(A,C)).

The mean percentage of subjects who generated the given conclusions for this category of task:

- Universal conclusions: 98.7%
- Mistakes: 1.28%
- Conclusions with "most": 1.28%
- Conclusions with "some": 0

[16] We give the proper numbers of syllogisms and not their position in the tests.

Empty-terms group (EG2): The results were similar, although slightly weaker, than for *CG*2. Again particular and "most"-conclusions almost did not occur at all (one particular conclusion for syl. 5). However we observed also a higher rate of mistakes in this group.

The mean percentage of subjects who generated the given conclusions for this category of tasks:

- Universal conclusions: 83.33%
- Mistakes: 15.27%
- Conclusions with "some": 1.39%
- Conclusions with "most": 0

The increased tendency to generate mistakes in this group can be explained by a slightly lower logic experience of the subjects with comparison to the control group.

The general observation is compatible with the results of the "semantics-checking" and "direct inferences" parts. If a conclusion with a universal quantifier is true, then this conclusion is preferred by subjects and it dominates possible conclusions with "most/mostnot" or "some/somenot". One way of explaining this effect could be to say that the latter conclusions require an assumption that domains are non-empty and that this assumption is just not made by subjects, who are hence perfectly logical and correct when they refuse this kind of existential import. We doubt however whether this is a proper explanation of this phenomenon, since then we would expect a significant difference between groups (EG2 and CG2). The alternative explanation would be to invoke the scalar implicature. If "most" and "some" implicate "not all", then subjects may ignore conclusions with these quantifiers in case they have already given a sentence with "all" as a response. Note, the the implicational scale assumes that inferences of the form *All*/*Most* and *All*/*Some* are valid, however it may be that they are disregarded for pragmatic reasons, at least when we consider active logical competence. A subject who has already given an informatively sufficient response may not even notice that another conclusion, which is included in the first one, would also be true. We are of the opinion that a detailed analysis of how scalar implicature suppresses inferring conclusions with "some" would enable a better understanding of what is in psychology vaguely called an "atmosphere effect".

"Most/mostnot"-conclusions

The variance of results for all possible tasks allowing "most/mostnot"-conclusions (cat. I & cat. III) was quite high: $v = 80.6$ (EG2) and $v = 56.2$ (CG2). For

Category	EG2(%)	CG2(%)
Category I & III	35.1	28
Category I	61.5	48.1
Category III	0	1.3

Table 2.10: Percentage of subjects generating "most"-conclusions

only category I it amounts to $v = 36.9$ (EG2) and $v = 27.6$ (CG2), and for III category (Aristotelian premises) these numbers were: $v = 0$ (EG2) and $v = 0.3$ (CG2) (scalar implicature at work! – the lack of such inferences for this category). This high variance in the case of category I may be an effect of difficulty differences between tasks. We got the best results of correct inferences in the case of syllogisms 1. and 2.: 65.38% of subjects in EG2 and 83.33% in CG2 generated correct "most"-inferences for each of these two syllogisms (identical results). These are both syllogisms of the I Figure – the first one with "most"-conclusion and the second one with "mostnot"-conclusion. (Neither of them uses "mostnot" in any of the premises.) The worst result in both groups we got for syllogism 7 (II Fig., "no" and "most" in premises): 37.5% (EG2) and 26.9% (CG2). Table 2.10 shows the mean percentage of subjects generating "most/mostnot"-conclusions for different categories of syllogisms.

Particular conclusions – so-called existential import

For particular conclusions (based on so-called existential import and not on scalar implicature of "most"), better results were obtained in CG2 than in EG2. Variance in the case of particular conclusions was also high – for all the tasks allowing particular conclusions it amounted to $v = 52.9$ in CG2 and $v = 23$ in EG2. Observation of variation of results for different kinds of tasks allows to assume that the following factors may influence an occurrence of a particular conclusion:

- occurrence of negation in the conclusion (*neg*)
- possibility of inferring also different conclusion: (a) universal, (b) "most" -conclusions (*clas*)
- possibility of implicature (*impl*)

The highest percentage of subjects who generated particular conclusions was obtained for category II restricted to only tasks allowing conclusions without negation. When other conclusions were also possible (including implicatures) or when the (particular) conclusion contained negation, the percentage of subjects generating particular conclusions decreased. Table 2.11 contains the mean percentages of subjects that generated particular conclusion for different categories of tasks. The between-group effect (EG2 versus CG2) was not significant. [17]

Category		EG2(%)	CG2 (%)
All syllogisms*	total	18	25.7
	positive	27	42.3
	negative	12.5	18.4
Category II	In total	31.9	45.5
	positive	**45.8**	**71.1**
	negative	25	32.7
Category III		1.3	0
Category I		6.25	15.2
Category V		37.5	51.9

Table 2.11: Particular conclusions

* – Category IV (only implicature-based conclusions) is not taken into account.

Any statistical estimation of the real influence of the above-mentioned factors (*neg*, *clas*, *impl*) was difficult because the numbers of tasks in the different categories were not equal – namely there were unequal numbers of tasks with negative and positive conclusions and also unequal numbers of tasks in each of the 5 categories of syllogisms. Moreover, because of the high deviation from normality (Shapiro-Wilk test) of our data, we could not rely on any parametric statistic in this case. ANOVA (which should be treated only as a hint) with the dependent variable *parcon* – number of particular conclusions generated for a given task in the whole group of subjects and *neg*, *clas*, *impl* (dummy-coded variables) as predictors was not significant. Only the factor *clas* (EG2: $F(1, 10) = 6.493$, $p = .038$ and CG2: $F(1, 10) = 7.593$, $p = .028$; the assumption of homogeneity of variances was satisfied in both cases) turned out to be a significant predictor of *parcon*. Thus the number of particular conclusions was significantly lower for the tasks which allowed also other conclusions (with the universal quantifier or with "most").

To check the effect of suppressing a particular conclusion by universal conclusions or conclusions with the quantifier "most", we conducted the following procedure. First we summed up the number of particular conclusions for each subject in each of three categories of tasks:

- category III (universal conclusion possible – 3 tasks) – *univ*

[17] The Mann-Whitney test was carried out to check the between-group effect in the number of particular conclusions generated for the tasks by subjects in EG2 and CG2. It was not significant.

- category I ("most"-conclusion possible – 4 tasks) – *most*
- category II & V (only particular conclusion possible or particular plus implicature – 6 tasks in total) – *part*

The outcome sums were weighted: divided by the number of tasks in the category. Thus, as the dependent variables we used the **mean scores** a subject obtained for each category. Friedman tests showed that in both groups there was a significant difference in subjects' mean score in each category. In the CG2: $\chi^2(2) = 27.23$, $p < .0001$; in the EG2: $\chi^2(2) = 22.3$, $p < .0001$.

The post hoc procedure was carried out to check which two of the three compared categories differed significantly. Wilcoxon tests were used for this purpose. After applying Bonferroni correction only p-values lower than .0167 are reported. In CG2 all three comparisons were significant:

Comparison *part – univ*: $Z = -4.137$, $T = 0$, $p < .0001$, effect size $r = -.574$.

Comparison *part – most*: $Z = -2.640$, $T = 51.5$, $p = .003$, effect size $r = -.366$.

Comparison *most – univ*: $Z = -3.017$, $T = 0$, $p < .0001$, effect size $r = -.418$.

In the EG2 the comparison *most – univ* was not significant, however the remaining two comparisons were significant:

Comparison *part – univ*: $Z = -3.453$, $T = 0$, $p < .0001$, effect size $r = -.498$.

Comparison *part – most*: $Z = -3.104$, $T = 11.5$, $p < .0001$, effect size $r = -.448$.

Thus, the mean number of particular conclusions generated by the subjects was significantly lower for those tasks that allowed also universal or "most"-conclusions than for those tasks that allowed only particular conclusions. A significant decline in the mean of generated particular conclusions was observed both when there was a possibility of inferring "most"-conclusions and of inferring universal conclusions. However, in the control group, the mean scores of *part* for category *most* and *univ* did not differ significantly.

Conclusion 3. *Particular conclusions were in a sense "suppressed" both by conclusions with "most" and by universal conclusions. Moreover, the frequency of generating particular conclusions was higher (but significantly only in EG2) for tasks allowing additionally only "most"-conclusions than for tasks allowing also universal conclusions.*

Implicatures

Subsequently we analyzed the category of tasks allowing implicatures as inferences (Cat.4 and Cat.5). The variance of results for all the possible tasks allowing implicatures was quite high in both groups: $v = 24$ in EG2 and $v = 33.53$ in CG2. According to our prediction, implicatures occurred slightly more frequently in the case of those syllogisms that allow only this kind of inference than in the case

Category	EG2(%)	CG2(%)
Cat. 4 & cat. 5 (all the tasks allowing implicatures)	25.92	21.37
Category 4 (tasks without classical inferences)	33.33	22.53
Implicatures without negation	48.6	46.15

Table 2.12: Implicatures in two-premises tasks

of syllogisms which allow also classical particular conclusions. Inferences with negation also seemed to occur less frequently than those without negation. (See Table 2.12.) We suspected that the occurrence of an implicature for a given set of premises should be dependent on the presence of negation (*neg*) and/or on the possibility of inferring also classical conclusions (*clas*). Thus classical conclusions should *suppress* the occurrence of implicatures.

The analysis of multiple regression with *impl* (number of implicatures generated for a given task in the whole group of subjects) as a dependent variable and two independent dummy-coded variables: *neg* and *clas* was not significant in EG2 and in CG2 it was difficult to interpret because of non-parametric data and small sample size. We were not able to draw any certain conclusions about the impact of the mentioned factors on the occurrence of implicature.

Note that higher frequency of implicatures was obtained in the case of empty-terms group than in the case of non-empty terms, which coincided with the lower percentage of "classical" particular conclusions in this first group. However the between-group effect was not significant.

Difficulty

Since we chose all the possible syllogisms which differ in any essential way, our examples varied with respect to syllogistic figures and quantifiers in premises. It is well-known that some syllogistic figures are more difficult for subjects than the others and that e.g. the use of negation should also make a task much harder. The influence of these factors is also reflected in the high variance of the results for different kinds of syllogisms. Still, the real impact of each of these factors is rather hard to estimate. An interesting observation is that syllogism 20 turned out to be especially difficult. In the empty-terms group a correct conclusion for this syllogism was not generated at all and in the control group it was given only once (3.8% of subjects). This is an extremely low result with comparison to the rest of syllogisms.

This effect may be connected with the specificity of the fourth figure, namely the fact that the major term (the predicate of the conclusion) is a subject of the major premise and the minor term (the subject of the conclusion) is a predicate of the minor premise.

Syl.20

$$\frac{\begin{array}{c} No(A,B) \\ Most(B,C) \end{array}}{Somenot(C,A)}$$

For tasks 19 and 18, which are also of fourth figure, the results were slightly better probably because in these cases the "classical" proper answer was "nothing can be inferred", whereas the implicature was quite easy to see in the case of 19 task ($41,6\%$ EG2 and 53% CG2) but somewhat harder in the case of 18 task (25% EG2 and 0% CG2).

Syl.19

$$\frac{\begin{array}{c} Mostnot(A,B) \\ All(B,C) \end{array}}{Some^{*}(C,A)}$$

The steps of reasoning are the following. One infers "Some A's are B" from "Most A's are not B" and then one gets a simple Aristotelian syllogism:

$$\frac{\begin{array}{c} Some(A,B) \\ All(B,C) \end{array}}{Some(A,C)}$$

the conclusion of which is equivalent to $Some(C,A)$.

Syl.18

$$\frac{\begin{array}{c} No(A,B) \\ Mostnot(B,C) \end{array}}{Some^{*}(C,A)}$$

Here we infer "Some B's are C" from "Most B's are not C" and then get an Aristotelian syllogism:

$$\frac{\begin{array}{c} No(A,B) \\ Some(B,C) \end{array}}{Somenot(C,A)}$$

Failures

We counted as failures those responses which were not classically valid nor were implicatures, nor were any non-typical but true inferences. For instance from premises:

No flying elephant knows Polish.

Most pteroklaki know polish.

one of the subjects concluded: *No elephant can speak Polish to a pteroklak.*

This conclusion is true in the light of the premises but still cannot be classified as any of the expected correct inferences. The same concerns the conclusion: *Some beings living on planet Venus are intelligent* inferred from the premises:

Most beings living on planet Venus are Venusians.

Most beings living on planet Venus are intelligent.

It is not a failure since it is a correct inference from the second premise, however it says nothing about the relation between Venusians and being intelligent. The same holds for: *Some German philosophers are not metaphysical idealists* inferred from:

Most German philosophers are metaphysical idealists.

Most German philosophers are not logicians.

We do not count the above as a failure since it can be treated as an implicature of the first premise, however again it says nothing about the relevant relation.

Among what we counted as failures, many of the responses can still be explained by a "hidden" and indirect influence of an implicature. It seems the strongest effect to have in the case of syl. 8 in EG2.

All talking frogs are enchanted queens.

Most living queens of Poland are not enchanted frogs.

In this case 11 out of 26 subjects (45.8%) (which is also 64.7% of all the failures for this syllogism) concluded: *Some talking frogs are living queens of Poland* (or equivalently: *Some living queens of Poland are talking frogs*), which is neither classically correct nor can be justified by the implicature of the second premise. Such inferences can, however, be justified by the implicature of the proper conclusion of these premises, which is *Most living queens of Poland are not talking frogs.* The doubtful point here is however that if this was the case, then we would expect both of these sentences (a proper conclusion and its implicature) to be inferred together, which happened only in 3 out of these 11 cases. Moreover, in CG2 only 5 out of 18 (27.77%) failures for this syllogism can be explained in a similar way. Hence, it is difficult to say if the above-described results are not an effect of some other cognitive mechanism.

Other observations

We observed, additionally, the following phenomena:

1. Although subjects were allowed to give multiple answers, they tended to generate only one conclusion. Double conclusions occurred very rarely. (More than 1 essentially different conclusion occurred 27 times for $480(24subj * 20tasks)$ given answers in EG, which equals 5.6% and 21 times for $520(26subj * 20tasks)$ in CG, which equals 4%.) If double conclusions occurred, they were usually equivalent or "almost equivalent" forms of the same conclusion, e.g. converse forms of "some"-sentences or converse forms of "no"-sentences.

2. Some problems with the grammar of double negation occurred frequently, namely "all" and "no" quantifiers were used as the same universal quantifier: "nie wszystkie" ("not all") as "żadne" ("no") or "żadne nie sa nie" ("no are not") as "wszystkie" ("all").

3. Subjects tended to use also restricted quantifiers, e.g.: "Some German philosophers who are metaphysical idealists are not logicians."

4. Modalities occurred as a method of lowering the certainty with which a conclusion was generated, e.g.: "Is it possible that...", especially (but not only) in case of implicature-based inferences.

2.7 Final conclusions

2.7.1 Main results

1. No difference in so-called existential import with respect to the sort of terms used in the tasks.

   Contrary to our predictions, the sort of terms used in tasks – whether they referred to some real objects or not – had no impact on subjects' reasoning. Neither the experiment with direct inferences nor the one with two-premises tasks revealed any significant between-group (empty-terms group versus control group) difference. The hypothesis that the ontological status of entities about which people reason has any influence on inferences $All/Some$ (resp. $No/somenot$) cannot be sustained. On the contrary, it seems that, when given sentences with empty terms, subjects just *presuppose* the existence of the alleged entities – what kind of existential presupposition it is we will see in the last chapter. Thus, it turns out that in natural language "Some A's are B" is a true sentence even if there are no A's nor B's (in the real world), since it says something about the relation between hypothetical A's and B's and nothing about their existence.

2. Significant effect of stronger so-called existential import of the quantifier "most" than of the universal quantifier. Scalar implicature of "some".

   According to our predictions subjects were more willing to infer particular sentences from premises with "most" than from universal premises. A conservative logician would probably explain this effect by saying that the inference *All*/*Some* is just not valid if no additional assumptions about non-empty domains are made. However such a hypothesis would not explain why subjects are not willing to give as a response a sentence with a quantifier "some" in case of a given (non-empty!) domain (given picture model) where both sentences are true: universal and particular. On the other hand, the weak scalar implicature of "some" perfectly explains the above result. Additionally the two-premise part confirmed this result. "Some"-sentences were not inferred in the case of syllogisms allowing universal conclusions. Thus we observed a *suppression* of particular conclusions by universal conclusions (as well as by conclusions with "most"). We suggest that this effect, known so far in psychology as the "atmosphere effect", should be rather explained with scalar implicature of the quantifier "some". The fact that conclusions with "some" were more strongly suppressed by universal conclusions than by conclusions with "most" is also coherent with the result that "most" has stronger so-called existential import.

3. The scalar implicature of "most" ("most" implicating "not all") was also checked to play an important role in reasoning.

   Arguments:

   (a) The experiment with picture-models confirmed that "most" has the assumed kind of scalar implicature. Sentences with this quantifier were not generated for models in which "all"-sentences were true (analogously for "most not").

   (b) In the part with direct inferences, "most"-sentences were the least willingly inferred conclusions form universal sentences. Furthermore, implicatures occurred as inferences from "most"-sentences ("somenot"-sentences for "most"-premises and "some"-sentences for "mostnot"-premises) as the most preferred responses.

   (c) In the two-premise part we also observed implicature-dependent responses. The variance was high and dependent on many factors but the mean percentage of subjects that generated implicatures for all the tasks which allowed this kind of inferences was not much different from the mean percentage of subjects who generated classical particular inferences for the relevant tasks.

4. "Some" as "not many".

As an additional result we got a plausible relationship between the cardinality of a set of objects to which we refer, or alternatively model-dependent cardinality (how big subset of a given domain is the considered set), and the probability that the particular quantifier "some" will be used by a subject. This special kind of scalar implicature of the quantifier "some" seems worth further and more detailed investigation.

2.7.2 Further research

The following further research seems worth consideration in the light of the obtained results:

1. An experiment oriented towards checking the cardinality dependence of "some".
2. The estimation of the difficulty of syllogisms with "most" – its dependence on the occurrence of negation in premises and syllogistic figure.
3. The domain- and context-dependence of the quantifier"most" – how it differs in bigger and smaller universes and how it depends on the the kind of universe. Additionally, it is worth checking how "most" behaves in universes logically divided in more than two parts. A picture-based experiment could help checking this property.
4. The occurrence of modalities in reasoning connected with implicature.

   Since subjects used frequently some modal expressions like "it is *possible* that... ", "some A's *may* be B", "not all A *have to* be B... etc. when it came to generating implicature-dependent conclusions, we consider this phenomenon worth investigating in the context of probabilistic or non-monotonic reasoning.
5. The divergence between active and passive competence with respect to the above phenomena.

   The most important next step should consist in an experiment oriented towards checking the above results with respect to passive logical competence. For instance, implicatures are very likely to work differently in those two cases. We mean here both implicatures connected with "most" and those connected with "some". We conducted such an experiment and it it described in the next chapter.

Chapter 3

Second experiment: passive competence

3.1 Layout of the experiment

3.1.1 Purposes

The second experiment was planned because of two main reasons. First of all we wanted to check with respect to passive logical competence some of the results we got in the first (active) experiment. We hoped, on one hand, to confirm some of our results, e.g. the difference between the so-called existential import of "most" and of "all", and on the other hand to look for plausible divergence between both mechanisms: active and passive. For instance we expected that passive evaluation of some inferences as correct (e.g. implicatures) would appear with much higher frequency than active generation of those inferences.

The second reason for conducting the passive experiment was the relatively weak statistical effect of the first one, which because of small sample sizes and the specificity of the layout (e.g. high variety of tasks in the two-premise part) has to be treated as half-qualitative pilot research. One of the main aims of the high diversity in our first experiment was to check what tendencies in replies for different kinds of tasks would appear, so that we could use the results later for formulating detailed hypotheses for statistical testing. Moreover, the high rate of responses that were difficult to classify and required prior interpretation was an additional reason to categorize the research more as a semi-structured interview rather than a purely quantitative study. Therefore we considered that a clear statistically significant result would be a good completion to our research.

The passive experiment was directed towards checking the following properties:

- *Scalar implicatures* of "most" and "some"; in the case of "some" – both the weak and the strong implicature (see chapter 1.3), in the case of "most" implicature in the particular and in the universal form.

- *So-called existential import* of "most" and of the universal quantifier, so the inferences: *Most/Some* and *All/Some* (resp. *Mostnot/Somenot* and *No/Somenot*).

  We have already shown in the first experiment that so-called existential import is not connected with any ontological commitment of quantifiers (the lack of between-group difference with respect to empty versus non-empty terms). It is proposed that these inferences are dependent on scalar properties and not on the alleged existential meaning of "most" or "some", and the lack of such meaning in the case of "all" (resp. "no").

- Inferences of the form *All/Most* and *No/Mostnot*, which we henceforth call "reversed implicatures".

The below working hypotheses were checked in the experiment:

Working hypotheses:

Hypothesis 5. *Scalar implicature of "most" will be evaluated as a correct inference.*[1]

Hypothesis 6. *The "weak" scalar implicature of "some" (Some/Somenot and Somenot/Some) will be evaluated as a correct inference.*

Hypothesis 7. *Inferences of the form Most/Some (Mostnot/Somenot) will display higher frequency than inferences All/Some(No/Somenot)(accordingly) (stronger so-called existential import of the quantifier "most" in comparison to the universal quantifier). We expect positive premises to display stronger existential import than negative premises.*

Hypothesis 8. *The "strong" scalar implicature of "some": Some/Mostnot and Somenot/Most will be evaluated as correct, but less frequently than the "weak" implicature.*

Hypothesis 9. *There may be a difference in the frequency of evaluation scalar implicature of "most" (resp. "mostnot") as correct depending on the form – particular or universal of this implicature. (For ease of exposition, we will further use the notions of "particular" and "universal" scalar implicature of "most" to name the corresponding forms.) At this point we do not, however, have a clear prediction concerning in which direction a form can affect this frequency.*

[1] Which means here and below that it will get "yes" replies. The subjects are not asked about correctness of inferences, but whether given conclusions are true in the light of premises. Nonetheless, a "yes" reply is interpreted as acceptance of correctness of an inference.

Hypothesis 10. *So-called "reversed implicatures" (All/Most and No/Mostnot) will be evaluated as correct inferences significantly less frequently than so-called existential import of the universal quantifier (All/Some resp. No/Somenot) but with no significant difference to so-called existential import of "most" (Most/Some resp. Mostnot/Somenot).*

Let us say a few words about the theoretical and empirical background of our hypotheses. In the previous experiment (two-premise part) we observed that particular conclusions were suppressed by universal conclusions and by conclusions with "most", however universal conclusions suppressed them much more strongly. We expected that the distance on the implicational scale may play a crucial role in suppressing inferences, so that *All/Most* inferences may be much more frequently recognized as correct than *All/Some* and also *Most/Some* much more frequently than *All/Some*, but there should bo no significant difference between *All/Most* and *Most/Some*. Thus, stronger so-called existential import of "most" in comparison to the universal quantifier should be confirmed but has to be explained by scalar factors, namely *Most/Some* will be more willingly inferred as a correct inference than *All/Some*, **since "most" and "some are closer on the implicational scale than "all" and "some"**. This is also connected with the expected lower frequency of the *Some/Mostnot* implicature – this latter fact having, however, deeper justification in the plausible vague semantics of both quantifiers: "most" and "some". If "some" means "at least two but not too many" (or "a small part of the domain"), then it is not clear whether a "small part" cannot be sometimes as big as "most", whereas it seems quite clear that it should not be as big as the whole set.

3.1.2 Procedure

The following experimental procedure was proposed. We conducted two separate sub-experiments checking the above:

- Test 1: so-called existential import of the quantifiers "most" and "all" (positive and negative premises) together with scalar implicatures of "some" – the strong and the weak.
- Test 2: scalar implicatures of "most" (the particular and the universal) together with inferences *All/Most* and *No/Mostnot* (reversed implicatures).

For both experiments we planned two independent groups, which may be treated as "basic" and "control". Thus there were 4 tests in total. All four tests consisted (each) of 60 direct inferences to evaluate. Subjects were asked if on the basis of a given sentence they can *utter* some other sentence:

Example:

You know that *most A's are B*.

Can you say on the basis of the above sentence that *some A's are B*?

YES / NO

A,B represent any terms used in questions. There was no reference to truth or correctness. Thus the questions checked *behavioral disposition* of subjects to utter or not given sentences if some other facts are known to them. There were 40 tested inferences in each test and 20 fillers. In all four tests we used as fillers inferences similar to the checked ones but logically *obviously incorrect*. By *obviously incorrect* we understand classically incorrect inferences, for which one cannot find any plausible explanation for correctness; they are not justified in terms of implicature or in terms of any existential load (or the lack of such load) of quantifiers. One example of such obviously incorrect inference is *Most/All*. It is hard to think of any model of reasoning in which such an inference would be regarded as correct. Few doubts may arise considering such a choice of experimental layout. The first is that there are only 20 fillers in the set of 60 tasks. However if we regard 20 plus 20 tested inferences, then it turns out that they work as fillers for each other. The other thing is that we do not use any real fillers – so tasks of a completely different form, that could distract subjects' attention. The reason is that first of all we did not want to make the tests too long, secondly that we had to instruct our subjects about what we expect from them, namely they were told that they should follow their intuition (language understanding, etc) while solving the tests. We consider that mixing our tasks with tasks of a completely different form would cause *too much* confusion and would make the experimental situation artificial.

Test 1 – so-called "existential import" & implicatures of "some"

Test 1 consisted of the following groups:

Test *1a* – the "most" group checking the so-called existential import of "most" and the *strong* scalar implicature of "some".

Layout:

- *Most/Some* (10) and *Mostnot/Somenot* (10)
- *Some/Mostnot* (10) and *Somenot/Most* (10).
- Fillers:
  - *Most/Mostnot*(5) and *Mostnot/Most* (5)
  - *Some/Somenot* (5) and *Somenot/Some* (5)

Test *1b* – the "all" group checking the so-called existential import of the universal quantifier, together with the *weak* implicature of "some".

Layout:

- *All/Some* (10) and *No/Somenot* (10)
- *Some/Somenot* (10) and *Somenot/Some* (10)
- Fillers:
  - *Some/All* (5) and *Somenot/No* (5)
  - *All/No* (5) and *No/All* (5)

We planed the following comparisons:

- between-subject:
  - *Most/Some* with *All/Some* and *Mostnot/Somenot* with *No/Somenot*.
  - *Some/Somenot* with *Some/Mostnot* and *Somenot/Some* with *Somenot/Most*.
- within-subject: with respect to the occurrence of negation, so : *Most/Some* with *Mostnot/Somenot* etc.

Test 2 – scalar implicature of "most"

Each of two groups checked a different form of a scalar implicature of "most":

Test *2a* – the particular scalar implicature: *Most/Somenot* (10) and *Mostnot/Some* (10)

Test *2b* – the universal implicature: *Most/notAll* (10) and *Mostnot/notNo* (10).

Both groups – *a* and *b* checked additionally the same reversed implicatures and use the same fillers:

- Reversed implicatures: *All/Most* (10) and *No/Mostnot* (10)
- Fillers
  - *Most/All* (5) and *Mostnot/No* (5)
  - *All/Mostnot* (5) and *No/Most* (5)

We planed the following comparisons:

- between-subject comparisons of the particular and the universal implicature.
- within-subject comparisons of scalar implicatures and reversed implicatures, positive and negative forms.

Comparisons between tests

Furthermore, we planed some comparisons between tests (thank to the unified set-up of materials, each of the four groups can be compared with the remaining three, provided that the number of subjects in each group is similar). The following comparisons were the most important:

1. Reversed implicatures with existential import of "all" (resp. "no") and with existential import of "most" ("mostnot")

2. Comparison of scalar implicature of "some" and "most"

"Yes" and "No" answers

Note that all tests consist of YES/NO tasks. The first version of the experiment assumed the possibility of giving a "don't know" answer. However, on the basis of a pilot research (c.a. 5 tests given to random people to solve), we observed that subjects did not tend to give this kind of reply. They chose only "yes" or "no" responses. Therefore we decided to exclude the "don't know" possibility for the sake of simplifying tests and future statistics. (If it happened that a very small percentage of subjects tended to give "don't know" replies, then it could be problematic for our statistical analysis.)

Choice of terms; materials

Based on the results from the previous experiment, which proved no significant difference in subjects' responses with respect to domains (empty versus non-empty terms), we decided to use in the second experiment *only empty terms of first category*, thus new-introduced, non-existing in language, terms. The main reason for our decision was to minimize contextual factor in the reasoning, so that people would not give answers based on their knowledge about world.

The following 10 pairs (an abstract term plus a property) were chosen and combined with different quantifiers in positive and negative sentences: *mermogliny* + pink, *buzaki* + green, *mroczniaki* + bring bad luck, *grombliny* + have claws, *mgłowce* + have a cap, *zarkotki* + have long ears, *trakloki* + intelligent, *wyszczyki* + two-coloured, *klawuchy* + have red tails, *leprokraki* + like cheese.

Each pair was used for each syntactic form only once, so that no identical task was repeated. In the case of tasks with only 5 trials (fillers), we chose 5 pairs of a term and a property and the remaining 5 pairs were used for another filler. The tasks in each of four types of tests were ordered randomly by a computer program. The English translations of these tests can be found Appendix C.

Groups

Each of the four tests was solved by an independent group of people. These were students aged between 19 and 39 (however only 6 people in total were over 30) from different, but mainly humanistic, faculties. The selection of people assumed that subjects should not have any logical experience – so not only mathematical or logical backgrounds, but also philosophical backgrounds were excluded. Psychology students were also not preferable. Subjects in each group were mixed and came from the following backgrounds: history, English philology, Russian philology, pedagogy, culture studies, criminal investigation techniques, law.

Table 3.1 compares group sizes and subjects' age.

GROUP	***1a***	***1b***	***2a***	***2b***
number of subjects	49	48	40	40
age average	22.63	21.16	23.5	22.67
age range	$19 - 32$	$19 - 32$	$20 - 39$	$20 - 33$

Table 3.1: Groups in the passive competence experiment

3.2 Results

3.2.1 General analysis of results

Existential import

So-called existential import was measured in tests *1a* and *1b*. The results for the quantifier "most" (test *1a*) were the following: 56.32% of all inferences of a form $Most/Some$ and 62.24% of all inferences $Mostnot/Somenot$ were evaluated as correct (so got "yes" replies). The average for this quantifier (both positive and negative) amounted to 59.28%. We connect this quite low result for the so-called existential import of "most" (this is a classically valid inference not requiring any additional assumptions about domains, so one might expect much higher or even close to 100% frequency of evaluation this inference as correct) with the strong scalar implicature of "some".

To compare, for the universal quantifier (test *1b*) such inferences were much less frequent. Only 38.75% of $All/Some$ inferences were evaluated as correct and merely 26.25% of $No/Somenot$ inferences (32.5% on average). Table 3.2 compares the corresponding numbers.

Implicatures of "some"

The strong implicature of "some " (test *1a*) was evaluated less frequently as true than the weak implicature (test *1b*). 57.96% of $Some/Mostnot$ and $53,67\%$ of

	most	*univ*
positive premise	56.32%	38.75%
negative premise	62.24%	26.25%

Table 3.2: So-called existential import (passive competence)

	some	*somenot*
strong	*mostnot* 57.96%	*most* 53.67%
weak	*somenot* 93.33%	*some* 92.3%

Table 3.3: Scalar implicatures for "some" and "somenot" (passive competence)

Somenot/Most inferences were given "yes" answers (average 55.82%), whereas for the weak implicature those numbers amounted to: 93.33% for *Some/Somenot* and 92.29% for *Somenot/Some* inferences (average 92.82%). The results are given in Table 3.3.

Scalar implicature of "most"

The particular scalar implicatures of "most" (test *2a*) amounted to 80.9% inferences evaluated on average as true. In this for inferences *Mostnot/Some* the average was 80.25% and for *Most/Somenot* it was 81.5%.

The average result for the same implicature in the form of a universal sentence ("not all" versus "not the case that no") was similar and amounted to 78.9%. However, the corresponding results for positive ("most") and negative ("mostnot") premises were divergent. As much as 85.25% of *Most/notAll* inferences were evaluated as true (so even more than for *Most/Somenot*) and only 72.5% of *Mostnot/notNo* inferences. This was a significantly lower score in relation to both the positive premise and also to the particular form of an implicature (see below analyses). This effect is probably a result of a complicated form of this task (complex construction with two negations). Results are pictured in Table 3.4.

"Reversed" implicatures

All/Most and *No/Mostnot* inferences we evaluated as true at the average level of 55.44%, in this: 57% of *All/Most* inference and 53.88% of *No/Mostnot* inferences were given "yes" replies. The results differed between tests (*2a* and *2b*), and were

	most	*mostnot*
particular	*somenot* 80.25%	*some* 81.5%
universal	*not all* 85.25%	*not no* 72.5%

Table 3.4: Scalar implicatures for "most" and "mostnot" (passive competence)

higher in test *2a*, the difference was however not significant. [Results for the test *2a*: *All/Most* – 64%, *No/Mostnot* – 60, 25%; for the test *2b*: *All/Most* – 50%, *No/Mostnot* – 47.5%.]

Failures

Between fillers we used obviously incorrect inferences. The percentage of failures for these tasks was not high, but it differed essentially depending on the kind of inference.

The highest rate of mistakes was observed for so-called "double-negation failures", so inferences *All/No* and *No/All*. 22.9% of *All/No* inferences were evaluated as true and 18.97% of *No/All* inferences, which gives an average of 20.83%. This confirms the result from the first experiment: that people tend to understand the quantifiers "all" and "no" in the same way (just as the universal quantifier)[2] and seem to have problems with double negation.

Quite a high rate of failures occurred also for the *Most/Mostnot* (19.18%) and *Most/Motnot* (15.91%) inferences (17.55% on average).

As far as inferring a universal sentence from a particular premise is concerned, mistakes were rather rare. Only 3.75% of *Some/All* inferences and 4.5% of *Somenot/No* interferences were evaluated as true (4.13% on average).

Universal sentence as conclusions from sentences with "most" (resp. "mostnot") were evaluated as true in 3.5% cases (the same result for *Most/All* and *Mostnot/No* inferences).

In case of incorrect inferring a sentence with "most" from a universal premise the percentages of failures amounted to: 5.25% for *No/Most* and 5.5% for *All/Mostnot* inferences (5.38% on average).

3.2.2 Analysis of distributions

All our data were not normally distributed which was checked with the exploratory analysis. Kolmogorow-Smirnov and Shapiro-Wilk tests showed a significant deviation from normality for most data in groups *1a* and *1b* and for all

[2]at least in Polish

dependent variables in groups *2 a* and *2 b*. Levene's tests showed that the assumption of homogeneity of variances was violated in most cases of group comparisons in test *1*, although it was satisfied for most comparisons in test *2*. Non-parametric procedures were chosen to check all our hypotheses.[3]

Furthermore, the visual analysis of histograms of frequency distribution allowed interesting observations. The data for reversed implicatures ($All/Most$ and $No/Mostnot$ inferences) displayed a clearly bimodal distribution with two maxima at values: 0 and 10. In this case it means that subjects tended to score 0 out of 10 for this kind of task, so replied coherently either "no" or "yes" to all questions of this form. This fact and the average result for this category oscillating around 50% shows that the considered inferences are highly dependent on the individual's understanding of the question in the task. Thus, people tend to have rather clear opinions about such inferences and are in these opinions divided half-half.

The data for scalar implicatures for "most" displayed a negative skew with a peak at 10. Thus most subjects scored 10 for this category, so replied "yes" to all questions of this form. This means that according to our expectations most subjects recognized such inferences as true and remained coherent while giving "yes" answers.

Data for so-called existential import of "most/mostnot" displayed a tendency to be asymmetric with a slight negative skew (with a peak at 10). Thus people in general tended to answer "yes", remaining coherent in this.

For the so-called existential import of universal quantifiers the distributions differed slightly between "all" and "no". The histogram for "all" was bimodal (with peaks at 0 and 10 – the left one slightly higher). The distribution of data for "no" was clearly positively skewed. The histogram displayed that almost half of people scored 0, so replied coherently "no".

The data for scalar implicatures of "some" had rather flat distribution in case of the strong implicature (which may be evidence of people's indecisiveness about such inferences) and displayed a strong right skew (again at the value of 10) for the weak implicature. Over 75% of subjects replied "yes" in case of the weak scalar implicature of "some" remaining coherent in their opinions.

3.2.3 Testing hypotheses

Detailed statistical analyses were conducted to test the hypotheses. We describe briefly the procedures, however a reader not interested in statistical details may skip this part and read only the conclusions under each subsection.

[3]SPSS 15.0 for Windows was used for all the analyses.

Hypothesis 5: Scalar implicature of "most" will be evaluated as true on a highly significant level.

χ^2 tests for scalar implicatures of "most" in both groups (*2a* and *2b*) showed a significant inequality of subjects' scores (ranging between 0 and 10)[4] for both positive and negative premises. (*2a* – for "most": $\chi^2(10) = 135.450$, $p = .000$ and for "mostnot": $\chi^2(10) = 140.950$, $p = .000$; *2b* – for "most": $\chi^2 = 73.3$, $p = .000$ and for "mostnot": $\chi^2 = 33.7$, $p = .000$)

Conclusion 4. *This result together with the analysis of histograms of frequency distribution for these variables displaying the clear negative skew in each of four cases (peaks at the value of* 10*) allows to conclude that **hypothesis 5 can be sustained**.*

Hypothesis 6: Weak scalar implicature of "some" will be evaluated as true on a highly significant level.

χ^2 tests for weak scalar implicature of "some" showed a significant inequality of subjects' scores $(0 - 10)$ for both positive and negative premises. For "some": $\chi^2(10) = 275.583$, $p = .000$ and for "somenot": $\chi^2(10) = 204.083$, $p = .000$.

Conclusion 5. *The above result together with the analysis of relevant histograms displaying the clear negative skew in both cases (peaks at* 10*) allows to conclude that **hypothesis 6 can be sustained**.*

Hypothesis 7: "Most" has stronger existential import than the universal quantifier. Negative premises have weaker existential import than positive premises.

Mann-Whitney tests were conducted to check the between-subject effect, namely to compare so-called existential import of premises with "most" with so-called existential import of universal premises. Positive and negative premises were compared separately, namely we compared "most" with "all" with regard to so-called existential import, and separately "mostnot" with "no". The analysis confirmed our predictions. Significantly more subjects evaluated conclusions with "some" as correct for premises with "most" ($Mdn = 6$)[5] than for premises with "all" ($Mdn = 2$): $U = 870$, $z = -2.235$, $p = .013$ (exact 1-tailed), effect size $r = -.227$. The effect was even stronger for the comparison between "mostnot" and "no". Significantly more subjects evaluated "somenot" conclusions as true for premises with "mostnot" ($Mdn = 6$) than for premises with "no" ($Mdn = 1$): $U = 478.5$ $z = -5.075$ $p = .000$, effect size: $r = -.515$.

[4]Subjects' scores for this variables ranged from 0 to 10 since there were as many as 10 tasks of a given form to evaluate.

[5]Range was 0 to 10

Further, we checked whether there is a significant within-subject difference with respect to negation, so we compared "most" with "mostnot" and "all" with "no" with regard to so-called existential import. We expected negation to weaken so-called existential import. Wilcoxon signed-rank test was conducted to check both within-subject effects. However, for the comparison between "most" and "mostnot", we got slightly stranger so-called existential import for "mostnot" than for "most" and the result was not significant even for the 2-tailed hypothesis. Only the comparison between "all" and "no" confirmed the directorial hypothesis. Subjects evaluated "somenot" conclusions as correct inferences from premises with "no" ($Mdn = 1$) with significantly lower frequency than "some" conclusions as correct inferences from premises with "all" ($Mdn = 2$): $z = -3.237$, ($T = 55.5$), $p < .0001$ (exact 1-tailed), effect size: $r = -.330$.

Conclusion 6. *On the basis of our analysis we conclude that, as far as passive logical competence is concerned, premises with "most" are characterized with stronger so-called existential import than universal premises. Negation weakens existential import, but only in the case of universal quantifier, so "no" has weaker existential import than "all", but "most"- and "mostnot"-premises do not differ in this respect. Thus* ***hypothesis 7 is sustained****.*

Hypothesis 8: *The strong scalar implicature of "some" is less frequent than the weak implicature.*

We expected the strong scalar implicature of "some" to be evaluated as a correct inference with much lower frequency than the weak implicature. Mann-Whitney tests were conducted to check the difference between the strong and the weak implicatures separately for positive and negative premises with "some". According to our predictions, significantly more subjects evaluated as a correct inference from positive premises with "some" the weak implicature ($Mdn = 10$) than the strong implicature ($Mdn = 6$): $U = 319.5$, $z = -6.478$, $p < .0001$, effect size:$r = -.658$. A similar result was obtained for premises with "somenot": the weak implicature ($Mdn = 10$) displayed significantly higher frequency than the strong implicature ($Mdn = 5$): $U = 283.5$, $z = -6.628$, $p < .0001$, effect size: $r = -.673$.

Additionally, we checked whether there is a within-subject difference with respect to negation in the premise. Wilcoxon signed-rank tests were conducted to compare frequency of the strong implicature of "some"- versus "somenot"-premises on the one hand, and frequency of the weak implicature of "some"-versus "somenot"-premises on the other. None of these tests were significant.

Conclusion 7. *Based on the above analysis we conclude that, as far as passive logical competence is concerned, the weak implicature of "some" is characterized by a much stronger impact on reasoning (measured as the frequency of evaluating by people such inferences as correct) than the strong implicature, so* ***hypothesis***

***8 is sustained**. Positive and negative premises with "some" do not differ in this respect.*

Hypothesis 9: Particular and universal scalar implicatures of "most" are accepted as correct with different frequency.

Mann-Whitney tests were conducted to check how a form of a conclusion (universal versus particular) influenced the frequency of accepting scalar implicatures of "most". Universal and particular forms of implicature were compared separately for positive and negative premises. There was no significant difference for the positive premise, but the result was *significant for the negative premise*. Significantly less subjects evaluated scalar implicature of "mostnot" as a true inference in the group in which it had the universal form "it is not the case, that no... " ($Mdn = 8$) than in the group where it had the particular form "some" ($Mdn = 10$): $U = 595$, $z = -2.090$, $p = .036$ (exact 2-tailed)[6] $r = -.233$. We explain this effect by the complicated form of this conclusion (negation used twice, or even *three* times – due to the double negation form of "no" in the Polish language).

Furthermore, the analysis was conducted to check the within-subject effect of negation, so we compared the frequency of "yes" responses to scalar implicature in the case of positive and negative premises with "most". This was done separately in tests a (particular implicature) and b(universal implicature). "Yes" responses were expected to occur less frequently in the case of negative premises. This effect was significant only in group *2b*. Subjects' frequency of evaluating as a true inference the universal implicature was lower in the case of a negative premise ($Mdn = 8$, form "not the case that no...") than in the case of a positive premise ($Mdn = 9$, form: "not all"). The Wilcoxon signed-rank test's result: $z = -2.602$, ($T = 98$), $p = .004$ (exact 1-tailed), $r = -.291$.

Conclusion 8. *We **cannot sustain hypothesis 9**. The universal form of implicature had a significant impact on the (lower) frequency of evaluation this implicature as a correct inference only in the case of a negative premise, however it is doubtful that it was indeed universality of this form that played the role. Rather we believe that it was a complex form of the considered sentence (three negations) that lowered subjects' willingness to evaluate it as true, especially that for the positive premise we did not observe any significant effect of the universal form. (What is even more, in the latter case the universal implicatures were even slightly* more *frequently evaluated as true than the particular ones.) Moreover, there was a significant difference between positive and negative premises with respect to subjects' evaluating the universal implicature as correct, which was not observed for the particular implicature.*

[6]Non-directorial hypothesis in this case

Hypothesis 10: *All/Most* (*No/mostnot*) **inferences will be more frequent than** *All/Some* (*No/Somenot*) **inferences, and similarly as frequent as** *Most/Some* (*Mostnot/Somenot*)**inferences.**

Since, to be correct, the inferences *All/Most* and *No/Mostnot* require the assumption of non-empty domains (due to classical definitions of these quantifiers), they can also be considered a case of existential import. We were interested in checking the difference between these inferences and the so-called existential import of the universal quantifier, thus *All/Some* and *No/Somenot* inferences. We expected significantly higher frequency of inferences *All/Most* (*No/Mostnot*) evaluated as correct than of *All/Some* (*No/Somenot*) inferences. Such an effect would confirm our theory that these inferences are dependent on scalar factors.

Mann-Whitney tests were conducted to check the between-group difference. We compared *All/Some* (*No/Somenot*) inferences from test *1b* with *All/Most* (*No/Mostnot*) inferences from test *2a* and separately with the ones from test *2b*, which served here as a control group.

The comparisons were done separately for positive ("all") and negative ("no") premises. Our hypothesis was fully confirmed for the comparison between 1*b* and 2*a*. In this case, significantly more subjects evaluated *All/Most* inferences ($Mdn = 9$) as correct than *All/Some* inferences ($Mdn = 2$): $U = 678.5$, $z = -2.433$, $p = .007$, $r = -.259$, and significantly more subjects evaluated *No/Mostnot* inferences ($Mdn = 8$) are correct than *No/Somenot* inferences ($Mdn = 1$): $U = 533$, $z = -3.650$, $p = .000$, $r = -.389$.

For the comparison between 1*b* and 2*b* the hypothesis was confirmed only for the premise with "no", namely in this case significantly more subjects evaluated *No/Mostnot* ($Mdn = 4.5$) inferences as true than *No/Somenot* inferences ($Mdn = 1$): $U = 724.5$, $z = -2.029$, $p = .021$, effect size: $r = -.216$.

Subsequently, we compared reversed implicature with so-called existential import of "most", thus with inferences: *All/Most* with *Most/Some*. This was done by Mann-Whitney tests separately for negative and positive premises. According to our predictions, there was no significant difference and based also on the mean percentages of evaluation such inferences as true we conclude that they have comparable frequency.

Conclusion 9. *We conclude that, in general, sentences with "most" are accepted as conclusions from universal premises more frequently than sentences with "some". More precisely, No/Mostnot inferences are evaluated as correct significantly more frequently than No/Somenot inferences, however a similar effect for the positive premise was not clearly confirmed. Since, what is more, inferences Most/Some and Mostnot/Somenot were evaluated as correct on an approximately similar level as the considered All/Some (No/Mostnot),* ***hypothesis 10 is partly confirmed.***

3.2.4 Additional analyses

Additional comparisons proved also a significant difference in frequency between "reversed implicatures" and scalar implicatures of "most" on the one hand and between scalar implicatures of "most" and the weak scalar implicature of "some" on the other.

Comparing implicatures of "most" with "reversed implicatures"

Scalar implicatures of "most" were accepted more frequently than "reversed implicatures". The result was significant for both particular (test $2a$) and universal (test $2b$) implicatures, positive as well as negative premises. Thus, subjects more frequently accepted "not all" implicature of sentences with "most" than they agreed to infer the latter from universal premises.

Wilcoxon signed-rank test results for group *2a* for positive premises: $T = 118$, $z = -1.716$, $p = .044$ $r = -.191$; medians: $Mdn_{most/somenot} = 10$ $Mdn_{all/most} = 9$. For negative premises: $T = 158$, $z = -1.992$, $p = .023$, $r = -.222$; medians: $Mdn_{mostnot/some} = 10$ $Mdn_{no/mostnot} = 8$.

The result was more clear in the case of group *2b*. For positive premises: $T = 46.5$, $z = -3.706$, $p = .000$ $r = -.414$; medians: $Mdn_{most/somenot} = 9$ $Mdn_{all/most} = 5.5$. For negative premises: $T = 151.5$, $z = -2.501$, $p = .006$, $r = -.279$; medians: $Mdn_{mostnot/some} = 8$ $Mdn_{no/mostnot} = 4.5$.

Comparing scalar implicatures of "some" and "most"

Finally, we compared implicatures of "most" with implicatures of "some". A non-parametric test for four independent groups was conducted to check the significance of difference between the strong scalar implicature of "some" (*strongsome*), the weak scalar implicature of "some" (*weaksome*), the particular implicature of "most" (*partmost*) and the universal implicature of "most" (*univmost*). The Kruskal-Wallis test was highly significant for both positive and negative premises: For positive premise (the dependent variable *scalar*) $H(3) = 50.629$, $p = .000$ (Asymp. Sig.) For negative premise (the dependent variable *negscalar*) $H(3) = 48.095$, $p = .000$ (Asymp. Sig.) This result means that there was a significant difference in frequency of evaluating as correct different types of implicatures: *strongsome*, *weaksome*, *partmost*, *univmost*, for both positive and negative types of premises taken separately.

A Post-hoc procedure (Mann-Whitney tests) was conducted to follow this finding. The following comparisons between implicatures of "some" and "most" were taken into account:

1. The strong implicature of "some" with the particular implicature of "most" (*strongsome* vs *partmost*)

2. The weak implicature of “some” with the particular implicature of “most” (*weaksome* vs *partmost*)

3. The strong implicature of “some” with the universal implicature of “most” (*strongsome* vs *univmost*)

4. The weak implicature of “some” with the universal implicature of “most” (*weaksome* vs *univmost*)

Bonferroni correction was applied so all results are reported at the level .0125. All comparisons except for the comparison: *weaksome* vs *partmost* were significant. P-values are reported for 2-tailed significance.

- Comparison 1 – *strongsome* vs *partmost*: The strong scalar implicature of “some” was significantly **less** frequently evaluated as true than the particular implicature of “most” for both positive and negative premises. The results were as follows – for *scalar*: $U = 469$, $z = -4.307$, $p = .000$; for *negscalar*: $U = 458$, $z = -4.393$, $p = .000$.

- Comparison 3 – *strongsome* vs *univmost*: The strong scalar implicature of “some” was significantly **less** frequently evaluated as true than the universal implicature of “most” for both positive and negative premises. The result for *scalar*: $U = 433.5$, $z = -4.566$, $p = .000$; for *negscalar*: $U = 621.5$, $z = -2.979$, $p = .003$.

- Comparison 4 – *weaksome* vs *univmost*: The weak implicature of “some” was significantly **more** frequently evaluated as true than the universal implicature of “most”. The results for *scalar*: $U = 640.5$, $z = -3.061$, $p = .002$; for *scalarneg*: $U = 548$, $z = -3.718$, $p = .000$.

Conclusion 10. *(Additional comparisons)*
Implicatures of “most” (in whatever form) are significnatly more frequently evaluated as correct inferences than $All/Most$ and $No/Mostnot$ inferences and more frequently than the strong scalar implicature of “some”. However the weak implicature of “some” is significantly more frequently evaluated as correct than the universal implicature of “most”.

Looking for correlations

Finally, we checked whether subjects’ responses to inferences with scalar implicature are negatively correlated with their responses to corresponding inferences of the weaker items from the stronger ones. This analysis was conducted to determine whether it is the *subject’s individual understanding of quantifiers that plays the role*. For example, assume that a subject understands “some” so that it is false in the case of “all”. In that case, to be consistent, she should reply “yes” for all *Some/Somenot* inferences and "no" for all *All/Some* inferences. Then we could conclude that "not all" is not implicature of "some" but a part of its meaning. However, if a subject accepts both inferences than the phenomenon requires a more complex explantation.

The following correlations where checked:

1. *Some/Somenot* inferences correlated with *All/Some*, and *Somenot/Some* with *No/Somenot*

2. *Most/Somenot* inferences (resp. *Most/Not..all*) correlated with *All/Most*, and *Mostnot/Some* (resp. *Mostnot/Not..no*) with *No/Mostnot*

3. *Some/Mostnot* inferences correlated with *Most/Some*, and *Somenot/Some* with *Mostnot/Somenot*.

The directorial hypothesis was assumed, namely we predicted a negative correlation in all cases. All the analyses but one (namely the third one and only for the positive premise) turned out to be not significant. "Yes" replies for *Some/Mostnot* inferences were negatively correlated with "yes" replies for *Most/Some* inferences. This means that the more the subjects were inclined to reply "yes" to one type of inferences, the less frequently they replied "yes" to the other. The correlation coefficient (Spearmans' rho) amounted to: $r = -0.373$, $p = .004$ (which is lower than the critical 0.01) ($N = 49$). As we mentioned above, there was no significant effect for the negative counterpart. This result could suggest that at least in the case of strong scalar implicature of "some", subjects evaluation of such inferences together with the corresponding *Most/Some* inferences is dependent on the *default definition* of "some" that is used by each individual subject. However, taking into account also the analysis of histograms for those variables, which are quite asymmetric or even flat in the *Some/Mostnot* case, we conclude rather that the result is accidental and dependent on subjects' indecisiveness about the considered inferences.

Conclusion 11. *We conclude that the mentioned inferences do not depend on individual differences in people's understanding of "some" or "most" (whether they use the standard definition or a definition with the implicature condition).*

3.3 Qualitative observations

Additional observations regarding qualitative aspects of our experiment were possible thank to direct interviewing one additional subject with the use of test tasks. We briefly describe an interesting observation based on comments given by the subject. The observation concerns how the content of tasks could have affected reasoning and caused differences between scalar implicatures of positive and negative premises. This concerns both implicatures of "some" (the strong and the weak) and of "most" (the particular). However, we analyze it for the example of "most".

Let us compare the following inferences:

$$(MN/S)\frac{Most(A,B)}{Somenot(A,B)} \quad (M/SN)\frac{Mostnot(A,B)}{Some(A,B)}$$

In the first case one states in the conclusion that some A's do not have a property B, whereas in the case of negative premise, the conclusion is positive and states that some A's *have* the given property B. This difference may be crucial, since whereas not having a property by an object does not imply any definite description of this object, the positive conclusion does imply such a definite (though partial) description. Let us analyze how it affects reasoning. In the below short dialogue we can clearly see that the subject's readiness to infer a conclusion is dependent on the term used as a predicate B. First we observe that the subject accepts the scalar implicature of "most".

Task A: *You know that:*

Most zarkotki do not have long ears

Can you say on the basis of the above sentence that some zarkotki have long ears

Subject[7]: *Yes, sure. If you say "most", then it is not all.*

However, later she refuses a "yes" reply based on the specific character of the predicate, although the logical form of the given inference is exactly the same.

Task B: *You know that:*

Most mermogliny are not pink

Can you say (...) that some mermogliny are pink?

Subject: *No*

Experimenter: *But just a while ago, you said "yes" for the identical question.*

Subject: *Hmmm, but it is not the same, it is totally different. With ears you can have either long or not long. But there are more colors than just pink.*

The problem arises since a property B is not of a "dichotomic" kind. There are probably properties which presuppose or suggest, or at least *allow* a dichotomous division of the domain. For example if it is not true that an object P is not long (like zarkotrytki's ears), namely if P is not *not long*, then we may conclude that

[7] female, 28, university background in tourism studies, profession: consulting company.

it is long. Thus the law of excluded law $(\neg B(P) \vee B(P))$ and also the rule of double negation elimination: $\neg\neg B(P) \vdash B(P)$ hold in this case. However suppose that it is not the case that P is *not* pink. Can we conclude that P is pink? In classical logic of course we can, since the two negations are eliminated, as in the previous case. Something may be pink or not pink, tertium non datur. However the subject does not infer such a conclusion and invokes as argument "many other possibilities". What kind of logical model may lie behind such a reasoning?

Let us analyze the considered inference (MN/S). By the assumption of the implicational scale that $Mostnot(A, B)$ is false if $No(A, B)$ is true and (by transposition rule) we have the implication $Mostnot(A, B) \rightarrow \neg No(A, B)$, with the consequent equivalent to $Some(A, B)$. But in fact our subject does not reason in that way. What she does is rather directly conclude form $Mostnot(A, B)$, which is equivalent to $Most(A, notB)$, that $Somenot(A, notB)$ (thus the mechanism is similar to the case of positive premise, where from $Most(A, B)$ one concludes that $Somenot(A, B)$). Then the conclusion $Somenot(A, notB)$ may or may not be reduced to $Some(A, B)$ by the law that is known in syllogistic as *obversion principle*. However, obversion seems to be rejected for some reason and although the subject concludes $Some(A, B)$ in the case of task A, she does not conclude $Some(A, B)$ in the case of task B.

Suppose then that the fact that P is not *not* B is transformed into a sentence that P is not *non-B*. Such a use of negation, where a term is formed from a term instead of a proposition being formed from a proposition is called *infinite* (Parsons, 2008). A term non-B is then true of exactly those things that are not B. From such a point of view a property non-pink is a *long disjunction*, namely non-pink things are those that are either blue, or green, or yellow, or red, etc. Now, if we know that P is not non-pink, we know that it is not blue *and* not yellow, *and* not red, etc, since according to De Morgan rules the negation of disjunction is a conjunction of negations. This forms a co-finite set, whose complement is a colour *pink*. Why is the conclusion then not inferred?

The reason may be that seeing the complement of such a co-finite set on a (potentially) infinite domain[8] is cognitively too difficult. The same process in the case of such terms as "long" is easy, since non-long is for most people just "short", and hence the complement of the set of things that are *not non-long* is just the complement of the set of things that are *short*. This point is also interesting, since it seems that people tend to treat properties called *opposites* ("long" and "short") as *negatives* ("long" and "non-long"). Of course things that are not long can be short, or medium, but such a trichotomy is not commonly noticed by people, who tend to categorize reality dichotomously. Yet with colors it is more difficult since pink is opposite to both green and blue, but blue is opposite to pink and green –

[8]The set of colors is quite large. Of course we know and are able to name usually only few colors, but we also know that there are many other possible shades, probably more that we can even think of. As it is hard to estimate the real number of colors, it is also hard to "grasp" the real mental representation of "non-pink"

the space of this category is not linear but has as many dimensions as there are colors (or at least a few dimensions depending on the number of basic colors). In this case a subject may be confused. She can treat non-pink as opposite to pink, and then take "other colors" as opposite to both non-pink and pink. Thus, contrary to other predicates, colors may be conceived as not satisfying the law of excluded middle – if something is not non-pink, it may be either pink, or it may have a complement color to this pair. The important element in the above-described cognitive process may be that people are representing objects rather as *having properties* than as not having them. Hence, an object that is just *non-B* lacks a representation, and to avoid this problem a subject may construct a mental representation of a negative property *non-B* with the use of a property opposite to B. If, however, *non-B* is a long disjunction then such a representation is problematic and can be substituted by some restriction of this set, e.g. for non-pink by "blue or green or yellow". If so, then of course not only pink, but also violet, red or orange are complements to non-pink.

The other (or complementary) reason for the above reply given by subject to task B may be the phenomenon recognized in psychology that applying De Morgan rules is difficult for people, e.g. $\neg(p \vee q)$ is often understood as $\neg p \vee \neg q$ (instead of $\neg p \wedge \neg q$). If such a mistake occurs, then the subject's reasoning is the following. She first concludes that some *mermogliny* are not non-pink, and then thinks that if so, then they are either not blue, *or* not orange *or* not red, etc... Then, of course, she cannot know what is the real colour of the remaining *some mermogliny*, since in each case there are many possibilities.

To what extend this model of reasoning is really applied cannot be estimated on the basis of our experiment. We would have first to analyze the terms used – whether or not they are treated "dichotomously", secondly we would have to compare subjects' readiness to infer $Some(A, B)$ in the case of the both kinds of properties. In our experiment we have not obtained any significant difference with respect to negation in the premise in any of the relevant types of experimental tasks. However, such a result is consistent with two possible models of reasoning:

- double negation is eliminated as in classical logic
- obversion rule is refused, however reduction is still done because in most cases the terms used suggest a dichotomous division of a domain.

The qualitative factors are not taken into account in our analyses and final conclusions. The modeling given in the subsequent sections concerns only the syntactic character of inferences and the quantitative results obtained in our experiments. We consider, however, this problem to be very interesting and worth future investigation. Interview-oriented experiments could allow to better judge how such inferences are processed and what kind of modeling should be applied.

3.4 Conclusions and discussion

Our hypotheses (except for hypothesis 9) were confirmed. The main results of the experiment concern:

1. Significant frequency of evaluating scalar implicatures both of "most" (over 80%) and "some" (over 90%) as correct inferences. Both χ^2 tests and analysis of histograms of frequency distribution confirmed hypotheses 5 and 6. Moreover, there was no significant difference between the weak implicature of "some" and the particular implicature of "most". However, a difference was observed for the universal implicature of "most" and the weak implicature of "some".

2. Significantly higher frequency of evaluating the so-called existential import of "most", i.e. inferences $Most/Some$ ($Mostnot/Somenot$), as correct than the so-called existential import of the universal quantifier, namely $All/Some$ ($No/Somenot$) inferences. A significant result was obtained for universality of the premise, so with respect to a quantifier in a premise: "most" versus "all" ("no"), and for the within-subject effect of negation, the latter however only in the case of universal premises.

3. Significantly higher frequency of evaluating as a correct inference the weak implicature of "some" than the strong implicature. Moreover the strong implicature of "some" was significantly less frequently given "yes" replies than the implicature of "most" (both in the particular and the universal form).

4. Additional comparisons displayed that "reversed implicatures" (average result 55.44%) were significantly less frequently accepted as correct inferences than scalar implicatures of "most" (c.a. 80%) (c.a. 1.5 times less frequently) and significantly more frequently than inferences $All/Some$ ($No/Somenot$) (32.5%) (c.a. 1.7 times more frequently). (Thus the distances were similar in both cases – c.a. $22 - 24$%.)

   "Reversed implicatures" were, however, approximately as frequently accepted as $Most/Some$($Mostnot/Somenot$) inferences (59.28%) and as frequently as the strong implicature of "some" (55.82%) (no significant difference between these three groups).

5. There was no significant negative correlation between subjects' scores for scalar implicatures and their scores for inferring the weaker items from the stronger ones, thus there was no correlation between $All/Some$ and $Some/Somenot$ inferences, $No/Somenot$ and $Somenot/Some$, nor between $All/Most$ (resp. $No/Mostnot$) and scalar implicatures of "most". The negative correlation was, admittedly, significant between $Some/Mostnot$ and

Most/Some inferences, but it was not significant for the negative counterpart, i.e. *Somenot/Most* correlated with *Mostnot/Somenot*.

The lack of correlation between the considered scores excludes the simple explanation that understanding of “some” resp. “most” is dependent on individual differences between subjects who either use these quantifiers with implicatures or without.

Let us say few words for those logicians who would rather claim that the low rate of so-called existential import of “all” and “no” is the result of incorrectness of such inferences in the case of empty domains (and actually all the used domains in our tasks were empty in the sense that they lacked reference in the real world). Such an explanation would not take into account the following facts:

- The results of our first experiment – that there was no difference in so-called existential import of “all” and “no” with respect to emptiness of domains.
- The high rate of *All/Most* and *No/Mostnot* inferences (in comparison to *All/Some* (*No/Somenot*)), which according to the standard account of existential import also requires an assumption that the domains are non-empty.
- The fact that such inferences (*All/Some* and *No/Somenot*) indeed were accepted as correct on the level of c.a. 30%, and by some subjects even consistently in all possible cases, so they were not excluded as incorrect (whereas *obviously incorrect* inferences were evaluated as correct on a much lower level and usually accidentally in single cases).

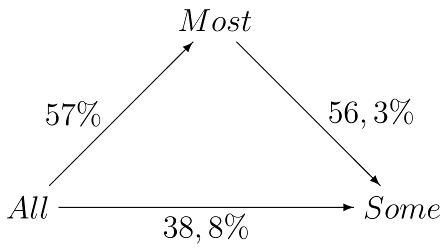
The last doubt with which we have to deal is that indeed the frequency of evaluating so-called existential import as correct was significantly lower in the case of “no” than in the case of “all”, which could confirm the theory that “no” in contrast to “all” lacks existential import. (See section 1.4.) Such an account would assume that “No A’s are B” means something like “It is absurd to assume that any A’s are B”.[9] We doubt, however, that it is existential import that matters here. Of course negative sentences behave slightly different than positive, but it is not connected with existential load. The interfering factor is the negation itself, which probably increases cognitive complexity. What is worth noting is that a Polish sentence “No A’s are B” uses double negation (in contrast to “Most A’s are B” or “Some A’s are not B”), and double negation was proved to be problematic for people. This might be the reason why there was a difference in so-called existential import for “no” and “all”, but there was no such difference between positive and negative premises with “most”.

[9]Experimental comparison of inferences *No/Somenot* with *No/Not..every* could shed light on this question.

3.5 Attempt of a theory: inflating quantifiers

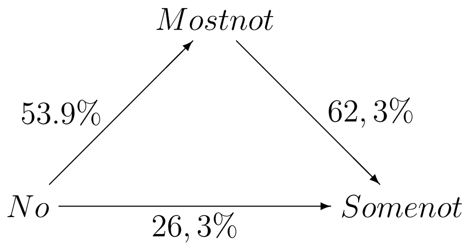
We believe that all the observed phenomena can be better explained in terms of scalar factors. Especially striking are the very strong effects of scalar implicatures of "most" (over 80%) and "some" (over 90%) (the weak implicature),together with the very low frequency of evaluating *All/Some* (*No/Some*) inferences as correct (ca. 30%). At the same time, *All/Most* (*No/Most..not*) occur with respect to frequency in the middle between the above two kinds of inferences (ca. 60%). The problem is that such a result only partly confirms the standard account of implicature and the relations between the items on the implicational scale. Let us recall that, according to the standard view on implicature, the use of the weaker item on the implicational scale implicates the negation of the stronger items, whereas all the weaker items are implied by the stronger. The implicational scale – the set of lexical items of the same category ordered according to their informativeness – assumes, in the case of quantifiers, at least the following orderliness: $\langle All, Most, Some\rangle$ (from the strongest to the weakest). Thus, although our experiment confirms the assumption that weaker items implicate the negation of stronger items (namely that both "some" and "most" implicate "not all"), it does not confirm the assumption that the stronger items imply the weaker ones. What is more, it seems that in the case of items from opposite poles of the scale ("all" and "some") this assumption is even contradicted, because *All/Some* (resp. *No/Somenot*) inferences are especially rare. While it is true that all used terms were "abstract" and empty, some previous experiments showed no significant difference with respect to emptiness of domains in subjects' inferring particular conclusions from universal premises. Since we refuse to connect the low frequency of accepting that kind of inferences with the existential load of "some" and "most", and the lack of such a load in the case of "all" (resp. "no"), we have to find another explanatory theory. Particularly interesting is the fact that, in the case of the stronger items implying the weaker, the frequency of evaluating such inferences as correct seems to reflect *distances on the implicational scale*. Thus *All/Most* is as frequent as *Most/Some*, and almost twice as frequent as *All/Some*. In the case of the weaker items implicating negation of the stronger items: both "most" and "some" implicate "some not" approximately with the same strength, however "some" implicates "most not" significantly less frequently. The strong implicature of "some" (*Some/Mostnot*) was significantly less frequent than both the weak implicature of "some" and the implicature of "most" and occurred on the level of ca. 60%, which again is not a clear confirmation of the assumed scale.

Summarizing, what is observed is a kind of non-compositionality of the semantics. The inference relation is not transitive any more: it turns out that even if $All \rightarrow Most$ and $Most \rightarrow Some$ hold, $All \rightarrow Some$ may not hold, as only half as many people accepted the latter inference than the remaining two.



If we look at the above triangle and the numbers assigned to each arrow, we can see that the inference relation *All/Some* is half as weak (in terms of statistical readiness of people to accept such inferences) as the component inferences: *All/Most* and *Most/Some*.

A similar triangle we get for negative counterparts.



3.5.1 Vagueness

First, let us observe that: "all", "most" and "some" are lexical items referring to a part of a given domain. Whereas "some" and "most" refer to their *denotations* vaguely, "all" represents the borderline case and thus is a sharp expression. Before we explain more formally what we mean by *quantifier denotations* and *vague quantifiers*, we describe roughly what it means for the quantifiers "most" and "some" to be vague and how they differ in this respect. We propose a distinction between "quantitative" and "qualitative" vagueness. The quantifier "some" is vague in the standard *quantitative* way, namely it is quite obvious that about 2 elements one can say "some", but it is not so obvious whether "some" refers equally well to 50 elements. This aspect of the vagueness of "some" is in general independent of the qualitative character of the domain, but it is probably dependent on the given or assumed cardinality of the domain, namely the meaning of "some" may not be connected with any *absolute* numbers but with the *relative size* of the referred subset of the whole domain. We do not exclude the qualitative vagueness of "some", however we believe that in each domain a *similar* kind of quantitative vagueness should be observed, and then the qualitative factor would play a secondary role. On the other hand, the vagueness of "most" is probably more of a qualitative character, namely whereas the quantitative meaning is sharp in most cases – "more than half" implicating "not all", the specific context (both qualitative character of the domain and the context of utterance) may involve specific understanding of majority.[10] Thus what we name here by "qualitative vagueness" is more a kind of context-dependence.

The below-described modeling does not refer to qualitative (for which purpose our experiment is not sufficient) but purely to quantitative vagueness, although we try to keep context-dependence in mind on each step. We assume the following:

- The default denotations of "some" and "most" can be extended.
- The default denotation of "some" is vague.

To explain what we mean by vague quantifiers and quantifier denotation we need to introduce some formal definitions. We recall first the alternative (to the one given in Chapter 1) definition of a generalized quantifier of type $\langle 1, 1\rangle$.

Definition 18. *A generalized quantifier Q of type $\langle 1, 1\rangle$ associates to each universe M a binary relation Q_M between subsets on M, s.t. if $(A, B) \in Q_M$, then A, B are subsets on M.*

Normally, it is assumed that generalized quantifiers are preserved by bijections, namely if $F : M \to M'$ is a bijection, then $(A, B) \in Q_M$ if and only if $(fA, fB) \in Q_{M'}$, for every subsets A, B of M, where $fA = \{f(x)|x \in A\}$ for $A \subseteq M$. Such a restriction, called *bijection closure*, follows from the assumption that quantifiers should be insensitive to which particular individuals happen to make up the universe, namely that *quanti*fiers are about *quantity* and not *quality* (they are not *quali*fiers). Note that if we want natural-language quantifiers to be context-dependent, then we have to abandon this condition.

Generalized quantifiers are usually defined in terms of a class of models – see Definition 8. Both definitions are, however, equivalent (assuming bijection closure in Definition 18). We chose this definition because we find it more clear and useful for our purposes.

Barwise & Cooper (1981) propose that all natural language determiners (i.e. in every language) are conservative.

Definition 19. *A generalized quantifier Q of type $\langle 1, 1\rangle$ is called conservative if and only if for all M and all $A, B \subseteq M$:*
$Q_M(A, B) \iff Q_M(A, A \cap B)$

Since in the present study we deal only with determiners, we restrict our investigations to conservative quantifiers.

Now, we use the notion of *witness set* introduced by Barwise and Cooper (1981) and define a witness set and the denotation of a $\langle 1, 1\rangle$ generalized quantifier *on a given domain* (first argument).

[10]Does it really suffice to find one over half stupid men to support a claim that "most men are stupid"? Probably not, since such a claim conveys a generalization the credibility of which increases with the percentage of men who are indeed stupid.

Definition 20. *A* witness set *of a $\langle 1,1\rangle$ generalized quantifier $Q(A,*)$, where $Q(A,*)$ means "a quantifier Q on a domain $A \subseteq M$", is any set $W \subseteq A$ such that $Q(A,W)$.*

The denotation *of a $\langle 1,1\rangle$ quantifier $Q(A,*)$ is the set of all witness sets of $Q(A,*)$.*

Note that:

Definition 21. The denotation *of a $\langle 1,1\rangle$ generalized quantifier Q associates with every M and $A \subseteq M$, the set of all witness sets of $Q(A,*)$.*

Thus, for example, a witness set for $All(woman,*)$ is a set of all women. In this case a singleton containing the set of all women is a denotation of this quantifier. With "some" it depends of course on how we define this quantifier. Assuming a standard definition of "some", a witness set for $Some(woman,*)$ would be e.g. a set {Cleopatra}, or a set {Cleopatra, Maria Skłodowska}, or even a set containing all women! The problem is that we have to reject the standard definition, since we claim that this quantifier is vague, which means that some witness sets of this quantifier are its witness sets only to a certain degree. Then, a set of all women is a witness set of $Some(woman,*)$ to a much lower degree than a set containing just 2 women. Therefore, the standard definitions are not sufficient. Below, we define the notion of a *vague quantifier*. For this purpose we need the notions of a *fuzzy* and *properly fuzzy* set, and a *membership ratio* function.

Definition 22. *We define a* fuzzy set A *as a function $A : \mathrm{dom}A \rightarrow (0,1]$.*[11] *We define a* membership ratio function *$in(e,A) = A(e)$ if $e \in domA$, and 0 otherwise. We say that a set A is* properly fuzzy *if there is an element e st. $0 < in(e,A) < 1$.*

Now we define fuzzy and vague quantifiers in terms of a fuzzy relation.

Definition 23. *A* fuzzy *quantifier Q of type $\langle 1,1\rangle$ associates with each universe M a fuzzy binary relation Q_M between subsets on M, s.t. if $(A,B) \in \mathrm{dom}Q_M$ (i.e. $in((A,B),Q_M) > 0$), then A,B are subsets on M. A relation is fuzzy if it is a* fuzzy set.

A vague *quantifier Q is a fuzzy quantifier, s.t. for some M, Q_M is properly fuzzy.*

The intuition behind a fuzzy set is that the higher the value of the membership ratio function for a given pair (e,A), the higher the degree to which the element

[11] We chose the set $(0,1]$ for the range since we wanted to avoid such problems as differentiation between a singleton with one element of a ratio 0 and the empty set. If we admitted elements with ratio 0 as elements of a fuzzy set, we would have to say how they differ from such objects that are no elements of this set with any ratio. In the case of quantifier denotations, we would have to determine why some sets are witness sets of a given quantifier to degree 0 and some are no witness sets of this quantifier.

e belongs to A. In the case of fuzzy quantifiers this results in higher truth-values for sentences with these quantifiers. The bigger r for $in((A,B),Q_M)$, the stronger the "truth" of $Q(A,B)$.

Notice that for a quantifier Q to be vague we only require that there is some model M such that Q_M is properly fuzzy. For example "some" in the model of cardinality 2 is not vague at all. One alternative way of defining a vague quantifier would be probably to demand a bit more, namely that Q is vague if it is a fuzzy quantifier, s.t. *for all sufficiently large models M (i.e. $|M| > n$ for some $n \in N$), Q_M is properly fuzzy.* Since we would rather avoid now a problem how big models will be large enough for e.g. "some", we continue with the weak formulation in which we require only that such a model M exists.

Since the above definition of conservativity cannot be applied to vague quantifiers, we define *fuzzy conservativity*

Definition 24. *A vague quantifier Q of type $\langle 1,1\rangle$ is called* fuzzy conservative *if and only if for all M and all $A, B \subseteq M$:*
$in((A,B),Q_M) = in((A,A\cap B),Q_M)$

We also have to reformulate the definition of witness set and explain what it means to be a witness set to a certain degree.

Definition 25. *We say that W is a* witness set *of a $\langle 1,1\rangle$ fuzzy quantifier $Q(A,*)$ to a degree r, if $W \subseteq A$ and $in((A,W),Q) = r$. The denotation of a fuzzy quantifier $Q(A,*)$ is a fuzzy set D such that for any W, $in(D,W) = r$ if and only if W is a witness set of $Q(A,*)$ to degree r.*

It follows directly from the definitions that:

Fact 1. *For every fuzzy $\langle 1,1\rangle$ quantifier Q, if Q is fuzzy conservative, then Q is vague if and only if for some model M and some set $A \subseteq M$, the denotation of $Q(A,*)$ is a properly fuzzy set.*

Proof. ($\leftarrow$) Assume there exists an A s.t. the denotation of $Q(A,*)$ is properly fuzzy. Take A, pick some W witness set of $Q(A,*)$ to degree r where $0 < r < 1$. Therefore, by the definition of denotation we have $in((A,W),Q_M) = r$, so Q_M is properly fuzzy, so by definition Q is vague.

($\rightarrow$) Assume Q - vague, then there exists M and exist $A, B \subseteq M$ s.t. $in((A,B),Q_M) = r$, with $0 < r < 1$, assuming fuzzy conservativity we have $in((A,A\cap B),Q_M) = r$, then since $A \cap B \subseteq A$, $A \cap B$ is a witness set for A to degree r. Therefore, there exists an A s.t. the denotation of $Q(A,*)$ is properly fuzzy.

□

Thus, a quantifier is vague if there are witness sets for this quantifier which belong to its denotation to a degree lower than 1. In the case of "some" this means that a given set of n blonde women can be a witness set for $some(women, blonde)$ just with some probability. Note that whereas in the case of generalized quantifiers we have minimal and maximal (with respect to inclusion) witness sets, these notions make no sense in the case of vague quantifiers.

The restriction of bijection-closure would be even less intuitive for vague quantifiers. Suppose Q is a vague quantifier and $f : M \to M'$ is a bijection. If we now assumed that Q is bijection-closed, we would have to demand that $in((A, B), Q_M) = in((fA, fB), Q_{M'})$, for any A, B of M, where $fA = \{f(x)|x \in A\}$. But such a condition seems to be a very strong restriction and we rather want to be liberal in defining the general notion of a vague quantifier. The lack of bijection-closure for vague quantifiers involves that there may be two different universes of the *same cardinality* such that some vague quantifier Q will behave differently in each of them, e.g. will be vague in one and not vague in the other, or the membership ratios will be different for witness sets of this quantifier in each domain. We do not want to exclude the possibility of such a context-dependent quantification.

Defining semantics for sentences with vague quantifiers requires the use of fuzzy logic, e.g. Łukasiewicz many-valued logic, with the set of real numbers as a set of truth-values. (Łukasiewicz, 1929, 1961) The fuzzy logic L, sufficient for expressing categorial sentences[12] with $\langle 1, 1\rangle$ vague quantifiers, can be defined in the following way. We take the language of classical first order logic, just instead of usual universal and existential quantifiers of type $\langle 1\rangle$ we have $\langle 1, 1\rangle$ fuzzy quantifiers. The formulas of our logic are built in a standard way, formulas with $\langle 1, 1\rangle$ fuzzy quantifiers are construed as formulas with $\langle 1, 1\rangle$ generalized quantifiers. Truth-value of atomic formulas is defined as in classical logic.[13] The truth-value for propositional connectives is defined as in Łukasiewicz. The truth-value for categorial sentence with a fuzzy quantifier Q is defined as follows:

Definition 26. *The set of truth-values is a set of real numbers* $[0, 1]$*,* F *– the set of formulas of the logic* L*.* $[.] : F \to [0, 1]$*.*

For any M*, any* v*-valuation, any* x*, any unary predicates* A, B*, any* $\langle 1, 1\rangle$ *fuzzy quantifier* Q*:*

$[Qx(A(x), B(x))]_{M,v} = in((A(x)^M, B(x)^M), Q_M)$

Since in our considerations we deal only with categorial sentences $Q(A, B)$, to simplify things we allow Q to be applied only to atomic formulas. Of course now we cannot say that "Some boys hate most girls", however such complex sentences are not presently of interest to us.

The important result is to show how we can get quantified sentences that are neither true nor false, but just "partly true" or "probable". One may assign the percentages on the triangle arms (see picture above) as weights of probabilities,

[12] sentences of a form $Q(A, B)$

[13] We define now the "minimal version" and omit the problem of fuzzy predicates.

with which such inferences are accepted by subjects. The other question is if one can treat them also as a kind of approximation of truth-values of such inferences.

Finally, let us briefly discuss another property worth considering with respect to vague quantifiers, namely *monotonicity*. In chapter 1.2.3 we define four different kinds of monotonicity for $\langle 1,1\rangle$ generalized quantifier Q – see Definition 15. Let us recall that $Most$, as a $\langle 1,1\rangle$ generalized quantifier, is monotone increasing in the second argument although it is not monotone in the first argument, whereas $Some$ is monotone increasing in both arguments. Defining monotonicity for vague quantifiers would, however, require deciding how the *ratios* (values of function in) change for extensions (resp. restrictions) of the domain (resp. the predicate), i.e. for subsets or supersets of the first or the second argument of the quantifier. To do this we would have first to determine how *in general* the ratios of membership of given witness sets to quantifier denotations depend on domains and their cardinality. As it is probably complex and dependent on many factors, we assume rather that vague quantifiers are *not monotone*.

3.5.2 Default denotations and inflating process

We have explained what we understand by vague quantifiers. Modeling the scalar implicature "not all" for both "most" and "some" requires however a little bit more work. Here we come to the essential point, namely we believe that quantifiers have *default denotations*, and the scalar implicature phenomenon (the weak implicature!) is the result of flexibility of those denotations, which can be dynamically changed in the process of reasoning.

Definition 27. *D is a* default denotation *of $Q(A,*)$ if it is a restricted set of witness sets such that for each $a \in D$, $in(a,D) \approx 1$ (We use $\approx$ instead of $=$ since we believe that the default denotation may be vague as well and may include elements which are* almost certainly *in the set.)*

The crucial issue is that in natural discourse we do not use sentences *to some extent*. We either use them or not, making a $0-1$ decisions whether or not they are true, depending on the context and dynamics of reasoning. Thus although there is some default meaning of language expressions, this meaning is *flexible*. The same concerns quantifier denotations which, at least in some cases of fuzzy quantifiers, can be "stretched" or extended. Note that the notion of default denotation makes sense both for vague and sharp quantifiers. In the case of sharp quantifiers such as "all" the default denotation is just inflexible.

What we propose is that the *default* denotations of "some" and "most" do not include the whole domain as a witness set for these quantifiers. More precisely:

- $Some(A,B)$ refers to the intersection of A and B of cardinality m, where $m < card(A)$

- $Most(A, B)$ refers to the intersection of A and B of cardinality m, where $m < card(A)$ and $m > card(A)/2$.

Moreover, the default denotation of "some" is vague, whereas the default denotation of "most" is flexible but not vague! This means that even if the use of the quantifier "some" in language suggests that the cardinality of any witness set should be smaller than *half of the domain*, this implicature is not ceratin. This is reflected in subjects' indecisiveness about inferences $Some/Mostnot$ (resp. $Somenot/Most$), which were accepted in approximately 60% of cases (the frequency distribution was asymmetric and displayed variance even for a single subject). Note that in the case of both quantifiers *the denotation itself* is vague, since in both cases it contains except for the default denotation the grey (or "possible") part, however in the case of "most" there is only one witness set outside the default denotation, namely the whole domain! We need to observe that vagueness which consists in flexibility of the default denotation differs in nature from what is usually understood by vagueness. We use the same notion because we apply fuzzy semantics for medeling this phenomenon. If one regards the case "all" as a possible witness set for "most", just lying outside the *default denotation* because of the very close to zero membership ratio, then indeed "most" may be treated as (*quantitatively*) vague. The difference between this kind of vagueness and the standard vagueness consists in the big truth-value gap between the default denotation and the remaining greyish part. Usually some regular (even *linear*) decrease of degrees to which elements belong to a vague set is assumed. In this case, we rather observe a "sudden decline". Moreover, ranking the elements from the grey part to the default denotation is a more dynamic process and involves increasing the membership ratios for those elements.

Now it also becomes clear why monotonicity of both "most" and "some" becomes problematic. Let us explain precisely what was already mentioned in the chapter 1.2.3 how the implicature "not all" (if embedded in semantics of quantifiers) forces us to abandon monotonicity. Let us consider that for some A, B, $Most(A, B)$ (with ratio 1) and assume B' s.t. $B \subseteq B'$ and $A \subseteq B'$. Then if $Most$ is monotone increasing in the second argument, we get $Most(A, B')$, but since $A \subseteq B'$, then also $All(A, B')$ and hence, assuming that the default denotation of "most" excludes "all", one gets $\neg Most(A, B')$ (or at least the ratio will be very close to 0). Similarly one can show that the vague quantifier "some" is not monotone in the second argument because of its implicature "not all". The non-monotonicity of "some" in the first argument follows from the vagueness of the default denotation. Assume for some A, B, $Some(A, B)$ to a degree r (i.e. $in((A, B), Some_M) = r$), and let A' s.t. $A \subseteq A'$, then assuming persistence of "some", we would have to claim probably that $Some(A', B)$ to degree r'. One could be tempted to say that $r' = r$, but as in this case r depends on the cardinality of A and of $A \cap B$, then since $card(A') > card(A)$, we would have to determine how r' will change in relation to r. We would like to say that $r' > r$, but the mutual dependencies are not that clear, since r can also depend on the qualitative character of A.

Finally, let us explain how the default denotations are stretched in the reasoning process. The most important observation is that, *implicitly*, both "most" and "some" do not refer to the whole domain. Thus, these quantifiers implicate "not all" ("somenot") and *are not used* if application of the quantifier "all" also results in a true sentence. Hence, a person who is given a sentence with one of these two quantifiers usually interprets this sentence as conveying also the considered implicature. The situation changes when people are asked to evaluate sentences with "some" or "most" as true or not in the light of universal premises. In such cases, $All/Most$ and $All/Some$ inferences can be understood in two ways. Asked whether on the basis of "All A's are B" one can say that "Some (resp. Most) A's are B", a subject may understand with or without the silent word "only" in the conclusion. Thus, the conclusion may be taken in the *default meaning* and hence rejected, or the meaning may be changed to "at least some (most)", and thus "some" or "most" will be understood without the implicature. The essential point is what cognitive mechanism lies behind the second interpretation. There may be two possible mechanisms:

- *suspension* of implicature,
- *extension* of a quantifier.

In the first case the implicature "not all" is suspended or canceled. In the latter we extend the default meaning of a quantifier, so that it embraces the grey part of its denotation. One might ask to what extend these processes are different, and why they are not regarded as the same mechanism. To see this, let us consider the quantifier "most". One possibility is that when we see (hear) a sentence with "all" we think (tacitly) also "most" – for example because of semantic accretion from voting systems (while voting we usually want *at least most* to agree for a given proposal, which thus is accepted also when all agree). Then, we may suspend implicature "not all" more easily in the case of "most" than in the case of "some". Still it does not seem clear enough why that kind of language habit connected with voting systems should have such a strong influence on semantics and hence why this cancelation of implicature should be easier in the case of "most". The other possibility is that the whole domain is treated as a kind of a "big majority" (thus it belongs to the grey part of the denotation of "most"), so we can extend the default denotation of "most" to the whole domain, whereas the same mechanism in the case of "some" is more far-fetched. Such extending denotations of quantifiers are possible because of their vagueness. The "grey" part of a denotation is like a non-inflated part of an elastic ball. If denotations are vague, they can be "stretched". The question is – how far.

Let us use an analogy of a two-colored (e.g. red and white) collection of little dots on a computer screen. "Some" refers to a situation where only a small part of the collection is red, “most” – to one where a bigger part is red and “all” to one where the whole collection is red. Thus a sentence “Some dots are red” assumes that *some dots are not red* and in most cases also that *a bigger part (most) of the collection is not red*. We are more hesitant about the latter conclusion because the default denotation of “some” is itself vague. On the other hand, “Most dots are red” means that a majority of the collection is red, but *not the whole collection*. Now, when we are given a sentence “All dots are red” and we are asked if sentences with “some” resp. “most” hold as well, we can either reject these sentences taking the default meaning of “some” and “most” into account or we can *abstract* in imagination from the whole picture an appropriate part to see that “some” and “most” can be special cases of “all”. This mechanism of abstraction will probably be more difficult if a picture is presented, but should be quite easy in the case of textually given models. The question is why such “abstraction” is easier in the case of “most” than in the case of “some”. If this is just suspension of implicature, it should be similarly difficult in both cases. That is why we consider that extending of the default denotations of quantifiers (“some” and “most”) plays a crucial role. This would explain why it is quite easy to extend “some” to “most” and “most” to “all”, but it is difficult to extend “some” to “all”. The reason is that the last case requires much more conceptual effort and is much more far-fetched, since the decrease of membership ratios of witness sets of “some” can be observed already when the cardinality of the witness sets exceeds half of the domain.

3.5.3 How to explain the hypotheses

Now we can finally clearly see how our model explains the hypotheses together with the results.

Hypotheses 5 and 6

Since the default denotations of “some” and “most” assume that both quantifiers do not apply to the case “all”, the scalar implicature “not all” is then the direct effect of such a semantics. Thus, the scalar implicature is a part of the *default meaning* of these quantifiers, although the default meaning may be extended in particular cases.

Hypotheses 7 and 10

“Some” may be easily extended so that the case “most” is treated as its witness set, but it cannot be easily extended to the case “all”, in other words the membership ratio for “all” being a witness set of “some” is very close to zero. That is why *Most/Some* inference are much more frequent than *All/Some* inferences. “Most” can be, however, extended to the “all” case much more easily than "some", because in the case of "most" the whole domain is the only witness set for this quantifier lying outside the default denotation. This results in *All/Most* inferences being more frequent than *All/Some* inferences.

Hypothesis 8

Inferences *Some/Mostnot* are less frequent than *Some/Somenot* inferences because of vagueness of the default denotation of "some", in other words, the case "most" belongs to the default denotation of "some" but with lower degree than e.g. the case "less than half", whilst the case "all" is outside the default denotation. This explains subjects uncertainty about *Some/Mostnot* inferences.

3.6 Summary and further research

We presented passive versions of experiments concerning scalar implicatures and the so-called existential import of the considered quantifiers. Based on our results we proposed how scalar implicatures of quantifiers "some" and "most" can be modeled with the use of fuzzy semantics. We describe these quantifiers as fuzzy with flexible default denotations. Those default denotations are further defined as not including the whole domain as an element. Therefore, the implicature "not all" is a part of the default meaning of both "some" and "most". Since the denotations are flexible, both considered quantifiers are extendable to the case "all", however such extension is more natural in the case of "most", since the whole domain is the only witness set for this quantifier lying outside the default denotation. In the case of "some" such extension is more difficult, since its default denotation is itself vague, which results also in people's indecisiveness about so-called strong scalar implicature of this quantifier.

As further research concerning vagueness and flexibility of quantifiers we plan the following:

- Estimation of the cardinality dependence of "some": approximating the borders of the default denotation
- Differentiation between the two mechanisms: suspension of implicature and extension of quantifiers

As an especially important completion to the above research we would consider the within-subject comparison of all the relevant inferences: *All/Some*, *All/Most*, *Some/Somenot*, *Some/Mostnot* both in the context (textual reasoning tasks) and with the use of pictures; for passive and active competence. Additionally, subjects should be asked to evaluate the certainty with which they infer their conclusions.

Chapter 4

What can fictitious discourse tell us about existence?

4.1 Introduction: what to do with empirical data?

This thesis began as a study of existential import, developed as a study of scalar implicatures of quantifiers and came back to the initial point. In our experiments we showed how all the inference relations between categorial sentences[1], including inferences that are traditionally said to invoke existential import (thus the inferences *All*/*Some* and $No/Somenot$), are dependent on scalar properties of the considered quantifiers. Moreover, we claimed that it is only scalar properties that matter here and that the considered quantifiers are free of any ontological commitment. Thus, the truth-value of categorial sentences does not depend on the ontological status of the entities referred to, so on the emptiness or non-emptiness of the domain. Whereas the first part of our claim was widely motivated on many pages of the previous chapters, the second one still needs analysis and justification. The problem yet remains and we are left with the question of how to interpret categorial sentences in predicate logic, since even if "some" has the suggested scalar implicature and vague denotation, it still has, in contrast to the universal quantifier, existential meaning.

Let us briefly recall from chapter 1, that the relationship between the ontological category of existence and the particular categorial sentence "Some A's are B's" results from how this sentence is represented in predicate calculus, namely as an existentially quantified sentence $\exists x(A(x) \wedge B(x))$. The further transfer of the existential commitment to the universal categorial sentence follows from what is traditionally understood by existential import, and is described as "subalternation" in the Aristotelian Square of Opposition. We mean here inferences:

[1] here – sentences of a form $Q(A, B)$, where $Q \in \{All, No, Some, Somenot, Most, Mostnot\}$

$$(A/S)\frac{All(A,B)}{Some(A,B)} \quad (N/SN)\frac{No(A,B)}{Somenot(A,B)}$$[2],

which in the present study are described as "so-called existential import".

However, since the universal categorial sentence is contemporarily represented as $\forall x(A(x) \rightarrow B(x))$ and thus can be vacuously satisfied (in contrast to the particular sentence), then if we want to preserve this kind of inference, we need to find some better logical model. [3]

The empirical evidence for the ontological independence of A/S and N/SN inferences consist in the lack of any between-group difference with respect to empty resp. non-empty domains. In our experiments the subjects' readiness to infer $Some(A, B)$ from $All(A, B)$ (or $Somenot(A, B)$ from $No(A, B)$) did not differ with respect to emptiness of domains. (Similarly for conclusions with "most" inferred from universal premises.) In general, even if we observed any between-group differences, they were of such a kind that cannot be connected with the presumed existential load of the quantifier "some" (and the lack of such a load in the case of the universal quantifier). Let us recall that in the direct inferences part of the first experiment, so-called existential import was even stronger in the empty-terms group than in the control group. This effect was explained by such factors as:

- differences in logical background of subjects in both groups,
- stronger context-dependence of reasoning with non-empty terms than with empty terms, namely that in the case of reasoning with the use of terms referring to entities that are familiar to us, it is difficult to abstract from factual knowledge about world.

Nonetheless, the experiments clearly showed that people's readiness to produce inferences of the form A/S or N/SN was not dependent on the ontological status of the objects referred to. This suggests that both quantifiers indeed have similar truth-value conditions in empty and non-empty domains. The question is, however, whether they are both existentially loaded (and hence both false in empty domains) or rather both free of any existential burden.

In this last chapter, we propose how to solve this puzzle. Let us introduce briefly what are our main assumptions:

1. Categorial sentences such as "All A's are B" or "Some A's are B" (among others) can be true or false independently of the ontological status of the

[2]Actually the latter inference (N/SN) in the Aristotelian Square of Opposition is not dependent on granting existential import to the negative universal categorial sentence (since "some not" lacks such import), but it requires additional assumption of non-emptiness of domains in modern logic.

[3]For a discussion of possible strategies of dealing with this problem – other than the one proposed below – see chapter 1.4.

domain. This assumption is the result of the following conviction. We believe that there is an essential difference in truth-value between the two sentences: "Zeus lived on Mount Olympus" and that "Zeus lived on Mount Tabor", namely the first one is true and the latter is false. Similarly "All elves have pointed ears" is in our opinion a true sentence, and not just because we believe that the universally quantified sentence can be vacuously satisfied. It is true since, indeed, according to what we know about elves their ears are pointed. On the contrary, a sentence "All elves have four legs" is clearly false, because elves are two-legged beings.

2. There is a crucial difference between so-called empty and non-empty terms. Let us make it clear. We do not claim that all terms are non-empty and hence that all objects are existent. Elves can be two-legged beings with pointed ears, and Zeus can be a Greek god, whose seat is Mount Olympus, but neither elves nor Zeus exist, at least not in the same world as ourselves. Therefore "Zeus" and "elves" are empty terms, namely terms non-referring to anything in the *real, actual world*. What is even more, elves inhabit different worlds than are ruled by Greek gods.

   It follows that "A's exists" is an ambiguous sentence, or rather it can have different presuppositions concerning the world or sets of worlds to which it applies. "Elves exist" is a false sentence if we are referring to the actual world (or if we refer to the world of Greek mythology), but it can be true, when understood as "In some fairy-tale worlds (e.g. Tolkien's world), i.e. some possible worlds, elves exist".

3. Existence indexes possible worlds, which are dynamic constructions.

   While under traditional approaches, possible worlds are treated as static objects and their internal structure is rarely examined, we take an approach in which possible worlds are dynamic mental constructions providing (possibly incomplete) descriptions of certain world-states (called *world-descriptions* or *worlds* from now). Neither objects nor worlds are primitive, thus worlds can be constructed of objects and objects can be also constructed within worlds if enforced by some world-laws defining them. Sentences about the real world are special only in the respect that the description of a world-state which is (often implicitly) conveyed in them is one of the real world (the world in which the sentence-using individual lives). Thus, existence is treated as a *modal index*, assigning objects to a certain world-state description. The most common natural-language use of the word "exists" is now treated as an implicit assignment of objects to the description of the real world.

In what follows, we develop each of these points, giving also some justification for our assumptions. Finally we will describe in more detail how our proposal solves the problem of existence and existential import.

4.2 How to separate existence from the existential quantifier

The problem arises when we try to represent natural-language sentences in the predicate calculus. If “Some A’s are B” is translated to “there are A’s that are B”, and hence if it is represented as $\exists x(A(x) \wedge B(x))$, then this sentence indeed states non-emptiness of some set, namely the intersection of A and B and thus it sates that *there are some objects*. It was Hume who said that there is no difference between existent objects and just objects, and hence that to be an object is to be an existent object. If so, then by saying that “Some A’s are B”, we are stating also that some objects *exist*.

One could, however, claim that to be an object and to be an existent object is not the same: that there are non-existent objects as well as existent. Within such an approach “there are A’s” is usually considered not equivalent to “A’s exist” and existence is claimed to be a normal first-order predicate like any other term naming a property, like “red”, “material”, “round”, etc. Then, the sentence “A’s exists” would be represented not as $\exists x A(x)$ (which would represent rather a sentence “there are A’s”) but with the use of a first-order “existence predicate” (Ex) and hence as: $\exists x(A(x) \wedge Ex(x))$.

In modern philosophy the possibility that existence could be a first-order predicate is usually rejected. The Kantian argument against “existence predicate” is invoked together with the famous quotation: “A hundred real thalers do not contain the least coin more than a hundred possible thalers.” (Kant, 2006) The idea of the argument is that if we have a full and adequate description of an object, we do not need to add existence as one of the object’s properties. Our description is already complete and stating existence will not give any new information. Thus, existence is different from normal first-order predicates. It is claimed to play a similar role to the truth-predicate, and hence is rather considered to be a second-order predicate. In those logics that stand in the Frege-Quine tradition, it is sustained that both “there is” and “exists” are expressed by means of the existential quantifier ($\exists$), which is, consequently, interpreted as having “ontological import”.

The view that granting existence to an object does not add anything to the description of this object is probably one of the greatest philosophical prejudices. Certainly adding to a description of an object P that P exists (or not) changes essentially our knowledge about this object. I can describe P and P may turn out to be my hallucination. If this fact is unknown to me and also to my interlocutor, then if she concludes that P exists – only because P has some properties which usually belong to existent objects – then her conclusion is false. Thus the information that P is not real, and so does not exist, is an essential completion to the description of P. In fact the qualitative descriptions of many fictitious objects, e.g. characters from “James Bond”, will not be much different from descriptions of real characters, even if evil characters from James Bond are much less dangerous to us than real bandits. Real and merely possible golden thalers in my pocket differ essentially, in that the latter are of no use.

At this point the defender of the view that existence is not a part of a description of an object might say that of course existent objects are different from non-existent ones and that she is not claiming anything opposite. Her point is that *descriptions* of those objects do not differ, since the information whether a description refers to an existent or a non-existent object is not a part of this description. These are properties used in the description that enable us to check if a described object exists, and using a property "existent" for determining whether it is really existent would be like reading the same newspaper twice to make sure that it is telling the truth (to use a popular philosophical metaphor). In other words, the information that an object exists does not change the conditions under which sentence asserting the existence of this object are true, and is not used in the procedure of verification. In this way objects are like sentences. True and false sentences differ a lot, but sentences "it is true that p" and "p" do not differ with respect to the truth conditions they have to fulfill to be true and the procedure of checking their truth-value is exactly the same. Moreover, saying that p is true will not make p true and similarly saying that P exists will not make P exist. The last argument appears to be a weak sophism in this respect that it works similarly with other properties: saying that P is red, will not make P red, since P is either red or not and, exactly the same way, P either exists or not.

Even if we find the above arguments weak, there is much more to say against the first-order "existence predicate". One of the strongest argument consists in showing how the assumption that existence is a (first-order) predicate allows constructing so-called *ontological proof*. We will not reconstruct this argument here (as it is well-known in philosophy and not relevant for our further investigations), but focus rather on so-called meinongean paradoxes of non-existent objects. If existence is a predicate which can be assigned to some objects or not, and thus we have both existent and non-existent objects, we need to answer the question "what does it mean *to be an object*".

In the most classical theory of non-existent objects, given by Meinong, to every single property and to every set of properties, there is a corresponding object, either an existent or a non-existent one. Thus, there is, for instance, an object that has the property of being blue as its sole property, one might call it "the object blue", or simply "blue". There is aslo an object that has the property of being round and the property of being blue, and no other properties (the object "round and blue"). And so forth. But "an object blue" cannot exist since it is impossible to have blue as a sole property. Certainly, "an object blue" is not identical with the property "blue", and it seems that every colored object needs to have also e.g. shape and size. Such an object ("blue") must be incomplete and indeed Meinong admits that such objects cannot exist, they are not only non-existent but even *necessarily non-existent* objects (Reicher, 2008). Furthermore, according to this theory, there can be also an object which is round and square (or even, to make it explicit, round and not round), which breaks the law of contradiction. Again, Meinong replies that the law of contradiction applies only to the existent objects and "round square" is a necessarily non-existent object. But this is not the end of problems. If "blue" is an object that has "blue" as its sole property, it has also a property "of having only one property", but then it has at least two properties, so we again get a contradiction.

Because of the above difficulties, non-existent objects together with existence predicate were not liked by many logicians and philosophers, although neo-meinongean theories, trying to find ways out of the aporias while preserving non-existent objects, also grew in large numbers. Later on we will refer to one of the most interesting, developed by Graham Priest, which uses a possible worlds semantics for this purpose. The reason why we do not want just to give up on these problematic non-existent objects is that identifying existence with the existential quantifier, or – using Quine's words – claiming that "to exist is to be a value of a bound variable", is not in the least deprived of weaknesses. Since we quantify in natural language over both real and merely possible objects, e.g. objects from fictional discourse, we are back to the problem of logical form and truth-value of sentences that use terms referring to such objects.

4.3 An existential journey into many worlds

Our proposal is a combination of intuitions coming from two different approaches to the problem of fictitious discourse. We mean the *intentional-story operator* and Graham Priest's application of possible worlds semantics to the problem of non-existent object. We below briefly present those approaches.

4.3.1 The intentional story-operator strategy

Let us use the already mentioned example about Zeus: From the two sentences: (a) "Zeus lived on Mount Olympus" and (b) "Zeus lived on Mount Tabor", the first one is obviously "truer" than the second one, since according to Greek mythology it was Olympus that was the seat of gods. Moreover, we deny the truth of the sentence (b) not because we do not believe in Zeus, but since it contradicts our knowledge about mythology, whilst (a) is clearly true in the light of this knowledge. This reference to context (mythology) leads us directly to the solution that sentences like (a) or (b), so treating about some fictitious entities, are in fact abbreviations of longer sentences with a silent so-called story-operator.

According to the story-operator strategy, we have to interpret sentences of fictional discourse as incomplete (Miller, 2008). A complete reading of, for instance, (a) *Zeus lived on Mount Olympus* would be as follows: (a') *According to the story S (Greek mythology): Zeus lived on Mount Olympus.*

The expression "according to the story S" is the so-called "story-operator". This is a sentence operator, so the whole sentence is in its scope. Thus (a) may be false if taken in isolation, but (a') the complete reading is true. What is crucial is that (a') does not imply that anyone lived on Olympus; neither does it imply that Zeus exists (or existed). The problem of this solution is that it does not work equally well for all kinds of sentences of fictional discourse. Consider, for instance: (d) *Zeus is a character from Greek mythology.* This sentence is true; but if we put a story-operator in front of it, we get a falsehood: (d') *According to the story S: Zeus is a character from Greek mythology*, since according to the Greek mythology, Zeus is an eternally living god (of flesh and blood). To avoid the above problem we may say that there are two kinds of sentences of fictional discourse: internal (e.g. (a)) and external (e.g.(d)) and the story-operator strategy can be applied to internal sentences only.

4.3.2 Non-existent objects and possible worlds

Graham Priest's (2005) strategy assumes that there are non-existent objects and existence is a first-order predicate. Then the Meinongean paradoxes can to be omitted (at least so is claimed by the author) with the use of possible words semantics.

The author introduces the distinction between *quantifier* and *ontological commitment*. Both the universal and existential quantifiers are treated in this approach as deprived of any existential import. A sentence $Some(A, B)$ has the form $\exists x(A(x) \wedge B(x))$, but the quantifier $\exists$ should be read just as "some" and not as "exists". Hence, $\exists x A(x)$ means just that some x are such that $A(x)$ and it does not mean that there are A's or that A's exist. Accordingly, by $Some(A, B)$ we just claim that *some objects from the domain are such that they have both: property A and property B* and not that such objects exist.

This theory gives "an object" a similar meaning to that given by Meinong. It is not just something real or existing, since there are non-existent objects as well. "An object" is anything that has at least one property. There are as many objects as there are combinations of properties, just not all of them are *existent objects*, namely not all of them exist in the *actual world.* Still, however ridiculous it may sound, *non-existent objects exist* too, but in other (non-actual) worlds, which is why they should be rather called "possible" than "non-existent", e.g. elves exist in the fantasy world, Zeus – in the world of Greek mythology. Moreover, like a real neo-meinongean solution, this proposal accepts *impossible objects* such as "round squares". They exist in *impossible worlds.*

Sentences from fictional discourse refer to non-existent objects and may be true or false just like sentences concerning real entities – with respect to the appropriate world in which they exist. Existence is a simple fist-order predicate which can be assigned to some objects in some possible worlds. "A chimera exists" is a sentence of a form $\exists x(Ch(x) \wedge Ex(x))$, where Ex is a simple first-order predicate that reads as "is existent". Quantification is ontologically neutral.

4.3.3 Modal index

Both above-described solutions are based on some good intuitions: the story-operator strategy – that sentences from fictional discourse refer to some given context, Priest's solution – that these "contexts" are possible worlds and that various objects exist in various worlds. The strategy we propose is quite similar to Priest's solution but avoids treating existence as a predicate.

What we propose is to move the existential import to the level of quantifying over possible worlds, so that existence becomes a kind of "modal index". Possible worlds are characterized qualitatively and existence of different entities in those worlds follows from the worlds' characteristics. For example fantasy worlds are such that they assume elves as their inhabitants, thus we know that in those worlds elves exist. Let us use here Kripke's example: if I dug out some remains of an animal similar to a horse with one horn, I would not discover that unicorns existed. What I would find out, would be rather that there lived some animals similar to mythological unicorns, but not that unicorns had ever existed, since unicorns are fictitious beings and one of their characteristics is that they do not occur in our world. Our world does not apply to unicorns, but a fantasy world does. On the other hand if I, in some way, opened a gate to another dimension and entered a different world with some fantasy flowers, trees and silver one-horned horses on the meadows, then I could say that I probably saw *real* unicorns (even if *real possible worlds* cannot be observed by any telescopes).[4]

Thus existence of some objects is their belonging to chosen possible worlds – a kind of function assigning objects to worlds. All the objects that belong to the actual world exist in the *basic* (actual or real) sense. Possible objects exists in the sense that they belong to some possible worlds (mythological, fantasy, worlds of imagination, etc.) Existence is hence ambiguous, but only in this sense that it indexes many possible worlds. Thus even if "to exist" is univocal, sentences of a form "A exists" are ambiguous, since they *presuppose* different models, in which their truth-value can be evaluated.

Sentences in natural language have *existential presuppositions* concerning in what kind of world (or set of worlds) we are, or which worlds we should consider to evaluate their truth-value. Sentences using so-called "non-empty terms" (so terms referring to some real objects) have a presupposition that we are in the actual world (we are considering the actual world's description). This is the *default existential presupposition* we make. On the other hand sentences from fictional discourse presuppose fictional world descriptions. That is why "Some elves have pointed ears" may be true even if there are no elves in our world,

[4] For this example and also for the first suggestion that existence could be a modal index – given in one of our philosophical conversations in Warsaw, I would like to thank to my friend Piotr Wilkin.

since this sentence is not about our world, but about some fantasy world (or worlds). This approach gives rise to multiple interpretations of fictional discourse sentences:

Ambiguity of presuppositions

Let $E = \{w \mid w \Vdash \exists x(elf(x))\}$ (Thus E is a set of all "elvish" worlds.)

1. (A) Some elves have pointed ears.
   - (A'-1) $\exists w : w \Vdash \exists x(elf(x) \wedge p.e.(x))$ [5]
   - (A'-2) $\exists w : w \Vdash \exists x(elf(x)) \wedge \forall x(elf(x) \rightarrow p.e.(x))$
     (Thus both $A' - 1$ and $A' - 2$ implicitly state that E is non-empty.)
   - (A") $w \Vdash \exists x(elf(x) \wedge p.e.(x))$, where w is a world-variable whose value is provided by the pragmatic context of the utterance.
2. (B) All elves have pointed ears.
   - (B') $\forall w \in E : w \Vdash \forall x(elf(x) \rightarrow p.e.(x)) \wedge E \neq \emptyset$ [6]
   - (B") $w : w \Vdash \exists x(elf(x)) \wedge \forall x(elf(x) \rightarrow p.e.(x))$, where w is a world-variable whose value is provided by the pragmatic context of the utterance.

In general, both A and B (below) have a similar presupposition – that we are in the set of worlds in which there are elves. The ambiguity lies deeper – when we have to deal with more than one possible world in which given objects exits.[7]

A sentence (A) "Some elves have pointed ears" may mean that (A') there is a possible world in which there are elves and elves in that world have pointed ears: (A'-1) some of them (but not all) or (A'-2) all of them. In both cases the sentence will be true provided that in case (A'-2) the scalar implicature of "some" is moved to the level of worlds, i.e. we assume that there are additionally some possible worlds in which elves exist, but do not have pointed ears. Thus it means that we know fairy-tales of elves with pointed ears, but also fairy tales of elves of different kind. This first reading seems most general and likely to appear if a sentence is given without any context.

In the second reading (A") we presuppose that we are in some chosen possible world (namely given by a context – explicit or presupposed on the basis of some pragmatic data, e.g. our interlocutor is a fan of Tolkien's books) in which there

[5] Abbreviation $p.e.$ represents a property "has pointed ears".

[6] This is where existential import is moved!

[7] We avoid now the problem of vagueness of "some" and assume for our analysis that a sentence (A) should be interpreted as "There are elves with pointed ears" with implicature "not all elves have pointed ears".

are elves and some of them (but not all) have pointed ears. Of course this interpretation requires the specification of a context, or may just appear if a subject is acquainted only with one kind of fairy-tale about elves (satisfying additionally this condition). In such a case, when we just choose one specific world, we of course do not have to care how many more possible worlds there are in which elves exist (with or without pointed ears), since everything is considered with respect to the given world. What is worth noting is that the scalar implicature of "some" has to work at least at one level: the level of worlds or the level of considered objects in the given world. It would not of course harm if we could apply scalar implicature on both levels and have many "elvish" worlds, some of them such that some (but not all) of their elvish inhabitants have pointed ears (A'-1).

As far as the universal sentence (B) "All elves have pointed ears" is concerned we again end up with two possible readings: one more general and context-free and the second more specific. Thus, (B) may mean: (B') in all possible fairy tales about elves (and *there are such*! – this is where the existential import of universal quantifier is moved to) all elves have pointed ears. This is however quite a strong requirement and as soon as the context of an utterance can be used to specify the possible world to which we refer, some subjects may understand this sentence as weaker and satisfied in less restricted conditions, namely that the universal condition must be satisfied only on the level of objects in the chosen possible world. Then (B) means that there is *the* possible world (the given one) in which elves exists and all elves in this world have pointed ears. The latter, although seemingly existential, can still be an interpretation of (B) even if there will be other worlds in which elves have normal ears. What is worth noting is that this is the point at which (A) and (B) *almost* collapse into one reading.

To understand our experimental situation it is also important to analyze, what happens when sentences use abstract, new terms, that are introduced into the language, e.g. "wyszczyki", "buzaki", etc., namely what kind of interpretation is given by a subject who is presented a task with such terms. Most probably such a sentence is then treated as *world-creating*. A subject assumes that "buzaki" or "wyszczyki" are just some entities, in whatever world they may exist. Such an assumption is not direct, but appears when the default presupposition about the actual world has to be suspended. The below dialogue pictures the reasoning process.[8]

> *Subject: (reading the first given question) "You know that some wyszczyki are two-colored..." What are wyszczyki? Is this a species of bird? Is this a tricky question?*
>
> *Experimenter: Not tricky. Do not focus on this, just answer the question, according to what you think.*

[8] A subject interviewed with the use of examples from the second experiment.

> *Subject: But this is not anything in the world, is it?*
>
> *Experimenter: No, it is not.*
>
> *Subject: Ok, well, anyway... I can conclude that some are not two-colored.*

What we can see is that, when the ontological status of objects in discourse is unknown to us, the default existential presupposition is made: that we are in the actual world. This presupposition may change when new information is given that this is just an invented term not denoting anything in the actual world. The default presupposition is suspended, but the whole sentence is not rejected as deprived of meaning (if we do not know what kind of entities are wyszczyki, how can we verify if some of them are red or not?) The sentence is instead taken as granted and its truth as given. The expression "you know that" is here essential, since the truth of the premise is ascertained. Thus, a subject is left with a problem of interpreting the given sentence – giving a model in which it can be true. It requires fixing the ontology, on which a subject can base her reasoning. Thus, a subject assumes a possible world description with such objects as wyszczyki, some of which are two-colored. A description is incomplete, but that is irrelevant since worlds, as we already mentioned are dynamic mental construction.

So-called existential import

Now we analyze so-called existential import assuming the above interpretations. First we observe that B' implies (if we leave out now the problem of implicatures) all three readings of A, namely: $A' - 1$, $A' - 2$ and A'', B'' implies $A' - 1$ and $A' - 2$, but A'' only if the world variable is the same in B and A. We suppose that it is the last interpretation (A'' resp. B'') that appears in the experimental situation, since premises are treated as world-creating, especially in the case of abstract (new in language) terms. Then the conclusion is generated (or evaluated) as referring to the the world-description given in the premise. It can of course be more complex in the case of empty terms referring to some well-known possible objects, as elves. Then the context may interfere in the process of establishing interpretation:

> *Subject [9]: "All elves have pointed ears..." Do you mean all possible elves I know, or some chosen kind of elves, like Tolkien's elves?*

In any case, such an assumption will lead to inferring so-called existential import, i.e. the sentence "Some elves have pointed ears", provided that this conclusion will not be understood as referring to some different world, which is not the one given by the premise.

[9] From the internet-forum discussion

4.3.4 Constructing worlds

According to the classical approach to modality, possible worlds are static structures, namely sets of objects with relations defined on them. One might wonder how a logical construct that is a possible world is actually constructed. Two main approaches can be proposed here:

- objects are treated as primitive
- laws are treated as primitive

In the latter approach, we have some laws that define a possible world and "force" the existence of some objects. In the former, we first select some objects and then just "throw" them into the world. We can, for example, create a world by selecting some real and fictitious objects and arranging them into a certain configuration; on the other hand, we can think of a possible world as an initially empty container that satisfies some global conditions, e.g. some physical or magical laws, and welcome into that world those very objects that are forced by these laws. To give an example of this approach: if in some world elves are like humans with respect to reproduction, then the existence of some elf in this world forces the existence of her parents. (In this case, we can also meaningfully speak of ontological commitment with respect to possible objects.) Those approaches need not, however, be viewed as mutually exclusive. In most cases, a possible world will be constructed *heterogeneously* with some primitive laws and some primitive objects. Possible worlds are hence *hybrid monsters* and the process of their construction is dynamic. This means that first we throw some objects from our world, postulating also some laws and, by their power, other objects are born into the world. One might think this approach faces a vicious circle problem. If some laws force the existence of a given object in a given world, does this object already exist in this world? If so, do the laws really force anything? The answer to this objection lies in the dynamics of the process of world construction. Possible worlds are not static. It makes little sense to speak of a world as a complete construction – what is given to us is just a world-description from the appropriate stage of the dynamic world-building process.

Another hybrid aspect of possible worlds is the possibility of having mixed worlds – possible worlds that are a mixture of the actual world and some fictional elements (or fictional world and some real elements). This possibility comes from allowing *individual objects* to "travel" between worlds. Of course, in most possible worlds there are some real objects, such as tress, animals, humans, etc. However, there is a difference between individual objects and the objects denoted by general terms. The latter are possible objects of a real species rather then real objects by themselves, and the individual designata of such species-terms do not have to come from the actual world. With objects denoted by individual terms the matter is altogether different. Let us consider a fable in which the author of this thesis discovers a magical gate to some fairy-land, where she meets elves. The possible world of such a story would be a hybrid of the real world and the "elvish" world. And more importantly, the main character in this world would be the real object (real individual human-being) from the actual world. In this sense, we believe in the *rigid designation* of individual terms. The identity of the author of this thesis is not changed in all possible worlds as far as she is referred to by her proper name.

Possible objects and paradoxes

First of all we observe that if existence is not a predicate but a modal index the problems that arise by using existence in the characteristics of an object (e.g. the ontological argument for existence of God) are avoided. What about the status of objects? Of course it might seem that we reject the Humean account that "to be an object" is "to be an existent object", but on the contrary we agree with this account, however we assume that existence has two meanings: the "real", when it indexes the actual world and the "possible", when it indexes some possible, but not actual, worlds. Thus "to be an object" is to be "existent object in some possible world". What is distinguished is hence not the existence itself, but the actual world.

The final questions concern how to cope with so-called "incomplete" ("the blue") and contradictory ("round squares") objects. In the case of contradictory objects we follow Priest's proposal that they can exist in those worlds in which the law of contradiction does not hold, so worlds with paraconsistent logics. Actually, we consider contradictory objects neither troublesome nor interesting enough to focus on them. Objects like "the blue" require, however, devoting few sentences to cope with them. We think that there are no such objects as "the blue". Not everything that can be characterized qualitatively, or that seems to be characterized qualitatively, is an object. One might ask, what is then "an object". The reply is that the notion of "an object" is primitive, which is why the statement that "to be an object is to be an existent object (in some possible world)" does not include a vicious circle. It is rather a claim than a definition. Defective objects are no objects, in no possible world in the universe of universes.

To see why defective objects like "blue" are rejected, we need to observe first that in a sense all possible objects are incomplete. This follows from the fact that they are, like whole worlds, dynamic constructs, or rather more precisely they are either real objects form the actual world or dynamic constructs. In both cases they are incomplete. The final complete descriptions of such objects are never available to us. What are available are only certain object-states i.e. descriptions of certain states of given objects. In the case of real objects incompleteness concerns how they will develop along the time dimension[10], whereas possible

[10] It is another philosophical question to reply whether past real objects, that existed and ceased to exist, are complete or whether they are more like constructs and hence incomplete.

objects may change in all possible dimensions, and hence may be incomplete also in *their* pastness. Let us take as an example Tolkien's elves. We know a lot about them as entities, how they looked, what kind of family relations they had, what was their history, what languages they spoke, and that they all sailed away to Valinor. But what about a sentence "Arwen wore a blue cloak on the day of her three-hundredth birthday" or "Arwen has a mole on the left side of her neck". The verity of such sentences may depend on other facts about the world she lives in, but most probably it is just a matter of the incompleteness of both the world and Arwen as an object. Then, both the world and Arwen may develop so that those sentences will be true or false. This leads us to even more interesting conclusions. Imagine that someone writes a new story set in Tolkien's world, with some new characters participating along with the ones already existing in the text. Do we want to say that it is still Tolkien's world or that it is a new possible world, let us say "Tolkien-plus"? Probably it would depend on how much the new author added to the original setting. From this point of view, possible worlds are fuzzy![11]

Thus, if all possible objects are incomplete, why do we reject objects like "blue". The problem is that Meinongean "incomplete objects" are not really incomplete, they pretend to be incomplete but in fact they are *ex definitione* complete. If "blue" has *one and only one* property, then it is a static, complete, though *defective quasi*-object, whereas possible objects may be incomplete constructs but they must be *potentially complete* and non-defective infinite collections of properties – like real objects, which they imitate.

4.3.5 Why not discourse semantics?

The last few words about why we did not decide to use discourse semantics for the analysis of existential import. Discourse representation theory (DRT), created by Hans Kamp (1981) is a framework offering a representation language for the examination of contextually dependent meaning in discourse. For example, anaphoric pronouns such as "he" and "she" rely upon previously introduced individual constants in order to have meaning. DRT uses variables for every individual constant in order to account for this problem. A discourse is represented in a discourse representation structure (DRS), which is a mental representation built up by the hearer as the discourse unfolds. The common semantic function of non-anaphoric noun phrases occurring in discourse is the introduction of a new

[11] One might ask what about the real world. This is difficult since as in the case of possible worlds, they are nothing more as mental constructs, in the case of the real world we must differentiate between the world itself and its representation, namely our ontology of the real world. For sure the latter (our ontology of the real world) is fuzzy. We would not stop conceiving our world of real just because one elf was born in the forests in the North. But it is counterfactual reasoning and shows fuzziness of the ontology and not of the world itself. To the metaphysical question whether the world itself is fuzzy we are not going to give any reply.

discourse referent (DR), which is in turn available for the binding of anaphoric expressions. DRS contains variables (DR) as well as sentences in formal language ordered like in the original discourse. Since discourse referents are mental representations the recipient has in mind the problem of ontological status of fictitious objects does not arise. Empty and non-empty terms are analyzed in the DRT in exactly the same way – as introducing new discourse referents.

Contrary to what may seem at the spot, our solution is to a large extent similar to DRT. Since objects as well as worlds are treated here in a sense as incomplete algorithms, mental constructs, one might understand worlds as discourse representation structures and possible objects as discourse referents. The reason why we engage possible worlds semantics rather than DRT is that in the latter approach the ontological problem is not touched at all and the philosophical problem of interpreting existential sentences about fictitious objects is left out unsolved. DRT provides only an algorithm for introducing discourse referents. At a next stage – when a truth definition of DRT is given – discourse representation structure may be embedded in "the world". However without such a mapping DRT remains at the level of mental representation and the question of reality of referents does not arise. What does it mean that some objects exist? This question is unanswered. On the other hand, the possible worlds semantics allows to deal with this problem in an elegant way.

Furthermore, possible world semantics allows to clearly differentiate between possible objects and their representations. The important point is that representations are always *representations of something* – of objects that exist independently, and they should not be identified with objects themselves. When I say that a horse is four-legged I do not mean that my representation of a horse is such, but a horse itself. The need of such a distinction is quite obvious in the case of non-empty terms. Why would we then mistake possible objects with their representations?

The reason is probably that possible objects are illusions. Of course there are no elves, no Greek gods, no flying horses and all sensible people know this. Thus, it seems that in this case we deal only with our images or so to say representations. However this is a misunderstanding. Zeus is not just my representation. I have of course a representation of Zeus, but Zeus himself is an *intersubjective* referent in a discourse. It is illusionary, incomplete, a kind of algorithm under permanent construction, but it is by no means my subjective representation. The same concerns possible worlds, they may be mental constructions, some may even be my own and private mental constructions, not known to anybody else, but as soon as they are shared in discourse by at least two individuals they obtain independent ontological status in the Fregean Third Kingdom.

4.4 Summary

In this chapter we proposed how to deal with the philosophical problem of existential import. We rejected both the possibility that the universal quantifier is characterized by existential import as well as the assumption that such an existential load is an element of the meaning of the natural-language quantifier "some". We suggest instead that quantification is ontologically neutral, and hence both sentences: with "all" and "some" may be true or false in domains that are traditionally considered as empty (thus with no elements from the real world), and hence that sentences describing real objects and sentences describing fictional or hypothetical objects share not only a syntactic, but also a semantic structure. The difference between such sentences lies only in the presupposed worlds to which these sentences refer, or rather the worlds in reference to which their truth-value may be determined. This kind of presupposition is called *existential* since it says in what kind of worlds the considered entities exist. The *default presupposition* in the case of so-called non-empty terms as well as new-introduced terms (whose world-affiliation is unknown) is that the actual world is the reference model. Sentences about existence are hence ambiguous. "Elves exist" is true when interpreted as "In some possible world there are elves" and false when read with the presupposition that the reference model is the real world. We propose that existence may be treated as a relation between certain world-state descriptions and object-states, both objects and worlds being dynamic constructs. It may then be represented as an index of a truth-value predicate, namely index referring to the model in which the truth-value of a natural-language sentence should be evaluated. Because worlds are treated as dynamic constructs, we discuss shortly the nature of the world-building process, which in most cases is supposed to be *heterogenous*, that is with some primitive world-laws as well as some primitive objects.

Chapter 5

Conclusions and perspectives

In our thesis, we discussed two related properties of reasoning with quantifiers in natural language: scalar implicatures and existential import. The close relation between these two phenomena arises from the analysis of the so-called implicational scale. Scalar implicature refers to inferring negations of stronger items from the weaker on the sale; whilst the stronger items are supposed to imply all the weaker. However, inferring both “most” and “some” from the strongest on the scale “all” requires an assumption that domains are non-empty or that the universal quantifier is, similarly like “some” and “most”, existentially loaded.

To investigate the above dependencies between quantifiers, we conducted reasoning experiments checking people’s active inference-production as well as their passive evaluation of given inferences. We observed that scalar implicatures occur with a high frequency in subjects reasoning. They are the most willingly generated inferences and also the most frequently evaluated as correct. On the other hand, weaker items are not willingly inferred from the stronger items, in this inferring “some” from “all” (resp. “somenot” from “no”) is especially rare. Moreover, the analysis of subject reasoning with the use of empty and non-empty domains showed no significant difference in people’s inferences with respect to ontological status of the objects referred to. Subjects inferences of sentences with “some” or “most” from universal sentences were invariant on whether fictional or real objects were mentioned.

Based on our results and especially on the lack of any connection between subject’s inferences and emptiness of domains, we explain all the observed phenomena in terms of scalar properties of quantifiers and refuse to analyze inferences with universal premises and particular conclusions in terms of existential import. We propose to treat quantifiers “some” and “most” as vague with flexible denotations. We provide a formalization of vagueness of quantifiers in terms of fuzzy semantics, and assume that quantifiers have default denotations which can be extended in the process of reasoning. Since default denotations of “‘some” and “most” do not include the whole domain as a witness set, the implicature "not all" is explained by the default meaning of these quantifiers. We also briefly discuss how the denotations are extended as well as how vagueness of the default denotation of "some" results with so-called strong scalar implicature of this quantifier.

Explaining all the inference relations between considered quantifiers in terms of scalar properties is not sufficient, since it does not solve the problem of interpreting existentially quantified sentences in fictional discourse. Therefore, to deal with this issue we propose to use a dynamic modification of possible worlds semantics, treating worlds as mental constructions describing certain world states. Existential sentences presuppose then a certain world or class of worlds that set up a domain of discourse. The most common natural-language use of the word "exists" is treated as an implicit assignment of objects to the description of the real world. What is more, in this approach not only worlds, but also objects are understood as incomplete dynamic constructs.

Our experimental study together with the proposed theoretical framework is by no means complete and should be understood as a part of a wider project. At the end of Chapter 2 and 3 we specified prospective research problems that seem worth considering in the light of obtained results. Summarizing the most important points we pose the following general open questions:

1. How "some" and "most" are used in small and big domains, how their semantics is dependent on cardinality and (presumably) on qualitative character of domains?

2. Are sentences with vague quantifiers treated as probable or partly true and if yes, then under what conditions people's certainty about them changes?

   Variants of the passive experiment with the use of a scale (both a number scale and a scale with modal expressions) on which subjects would estimate how certain they are about the inferences they accept could shed a light on this question.

3. Are there any individual differences in the way people understand "some" or "most" and make inferences with these quantifiers? Thus, are implicatures somehow dependent on individual understanding of quantifiers?

   Although our present experiments deny such an option, it should be still carefully analyzed and confirmed with further research.

4. Which of the cognitive mechanisms is more plausible to lie behind the inferences $All/Some$ (or $All/Most$): suspension of implicature or extension of quantifiers?

5. Finally, in relation to the above, how do active and passive competencies differ with respect to considered inferences?

Appendix A

Regression analyses from Chapter 2.4.3

Linear regression

For the linear **regression of *some* on *bdot***, ANOVA was not significant ($F(1,14) = 4,485$; $t(14) = -2,118$; $p = .053$). Moreover, the assumptions of analysis were violated. Scatterplot (zpred vs zresid) displayed some heteroscedasticity and the assumption of normally distributed residuals was not satisfied.

The three remaining analyses were significant, however they were difficult to interpret because the assumptions were again violated.

Regression of *sn* on *proc*:

R was significant for regression ($R^2 = .707$) (adjusted $R^2 = .686$), which means that 70, 7% of variance in the dependent variable can be predicted by scores of the independent variable. $F(1,15) = 33,805$; t-test for coefficient ($B = .153$) $t(14) = 5,814$; $p < .0001$. (Constant $B_0 = 9,097$ $p < .0001$) The assumption of normally distributed residuals was violated! The other assumptions were satisfied.

Regression of *sn* on *bdot*:

R was again significant ($R^2 = .64$) (adjusted $R^2 = .614$). $F(1,14) = 24,884$; t-test for coefficient ($B = 1,441$) $t(14) = 4,988$; $p < .0001$ (constant $B_0 = 9,968$ $p < .0001$)

The assumption of normally distribution of residuals was again violated, also the plot of regression standardized residual revealed some non-linearity. The Durbin-Wason test had slightly too law value (1, 557).

Regression of *some* on *proc*

R for regression was significant but low. $R^2 = .318$ (adjusted $R^2 = .269$). $F(1,14) = 6,514$; t-test for coefficient ($B = -.121$) $t(14) = -2,552$; $p = .023$ (constant $B_0 = 22,967$ $p < .0001$) Again the residuals were not normally distributed and scatterplot (zpred vs zresid) revealed some heteroscedasticity.

Cubic regression

ANOVA for cubic regression of ***some* on *bdot***: $R^2 = .612$ (adjusted $R^2 = .515$); $F(3, 12) = 6.309$; $p = .008$.

T-tests for coefficients for *bdot*: $t(12) = 3.115$; $bdot^2$: $t(12) = -3.269$ and $bdot^3$: $t(12) = 3.119$ were all significant on the level of $p < .01$ (Constant was not significant.)

The model equation:

$$some = 4.107 + 13.614bdot - 3.218bdot^2 + .202bdot^3. \tag{A.1}$$

ANOVA for cubic regression of ***some* on *proc***: $R^2 = .863$ (adjusted: $R^2 = .829$); $F(3, 12) = 25.224$; $p < .0001$

T-tests for coefficients for *proc*: $t(12) = 6.194$, $proc^2$: $t(12) = -6.132$ and $proc^3$: $t(12) = 5.731$ were significant on the level $p < .0001$ (Constant was not significant.)

The model equation:

$$some = -4.544 + 2.172proc - .052proc^2 + 0proc^3. \tag{A.2}$$

Logarithmic regression

Logarithmic regression of ***sn* on *bdot***: $R^2 = .744$ (adjusted: $R^2 = .726$); $F(1, 14) = 40.745$, $p < .0001$. T-test for coefficient: ($B = 5.712$) $t(14) = 6.383$, $p < .0001$. Constant $B_0 = 8.497$ was significant: $p < .0001$.

Logarithmic regression of ***sn* on *proc***: $R^2 = .777$ (adjusted: $R^2 = .761$); $F(1, 14) = 48.832$, $p < .0001$. T-tests for coefficient: $B = 6.106$, $t(14) = 6.988$, $p < .0001$. Constant $B_0 = -6.569$ was not significant.

Appendix B

Tests from the first experiment

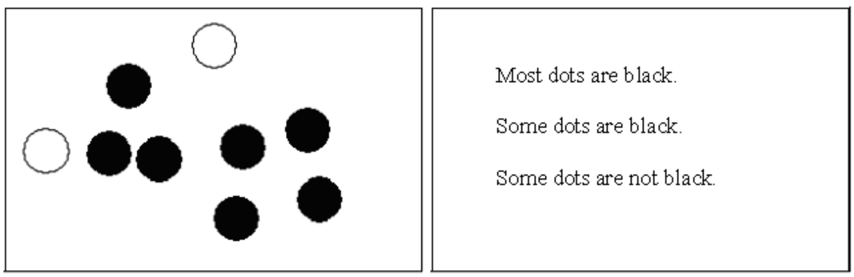
Picture test:

Gender:

Age:

Field of study (level):

Tutorial:



Looking at the above picture you can say that:

1. Most dots are black
2. Some dots are black
3. Some dots are not black

Since these sentences are true about this picture.

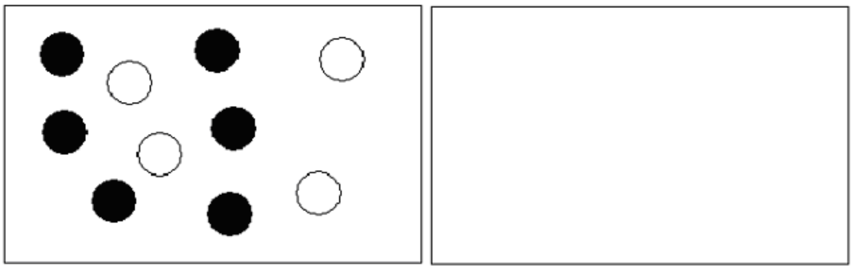
You cannot, however, say that:

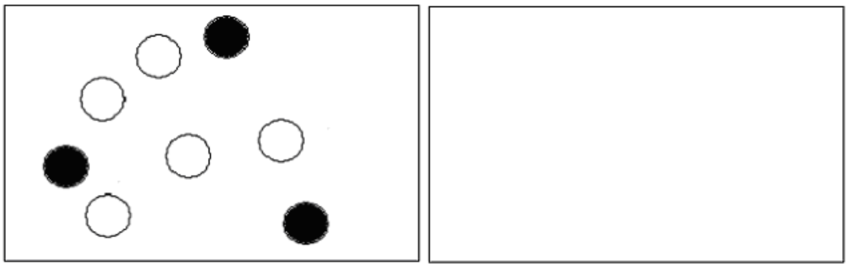
1. All dots are black
2. Most dots are not black
3. No dots are black

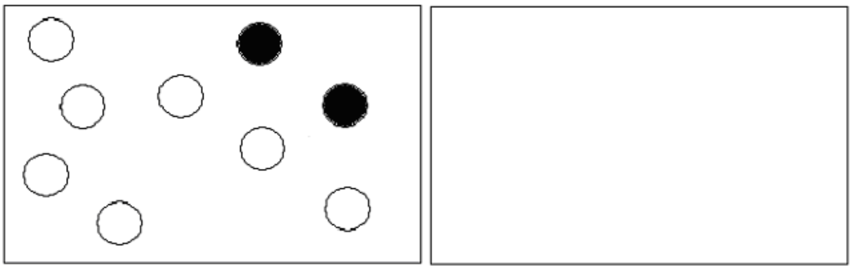
Since these sentences are false about the picture.

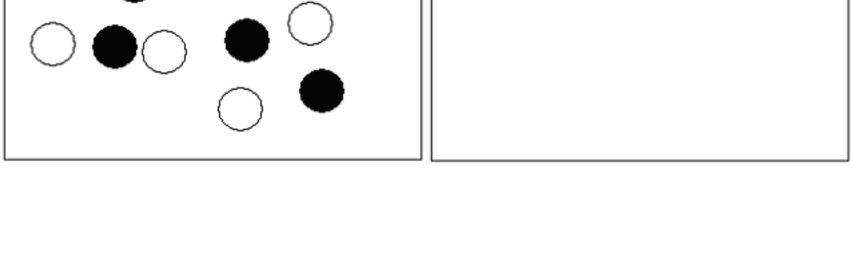
Instructions:

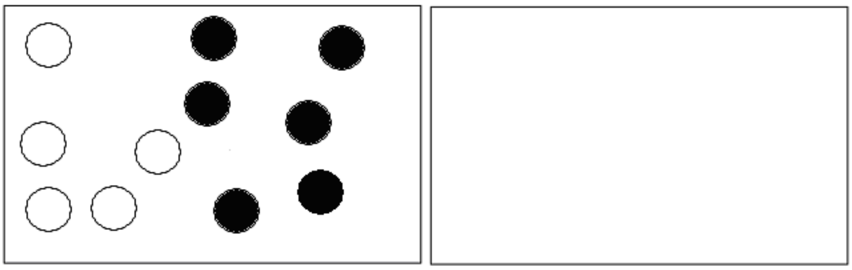
Look carefully at the below pictures. There are 20 of them. In the box next to the picture write down all the sentences (of form **a,b,c,d,e** or **f**) you think are true about the picture. The sentences should begin with words: **"most"**, **"all"**, **"some"** or **"no"** and can be both **positive** and **negative**. Please note that there can be **more than one** true sentence. Try to write down all the sentences you think are true. If you think you that no sentence is true, draw a **line**.

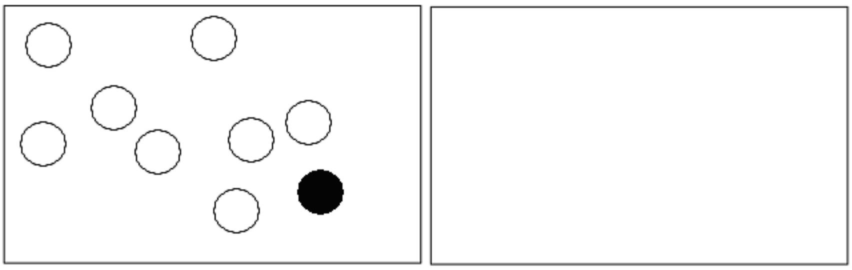


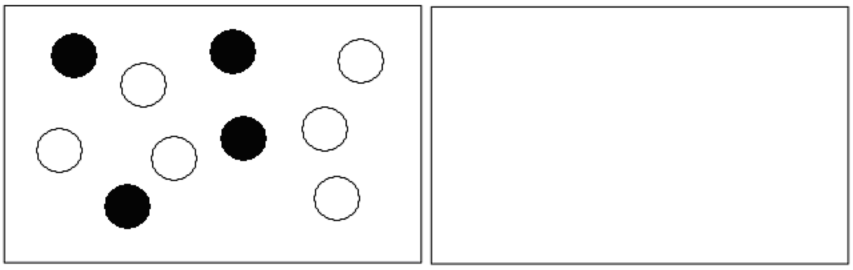


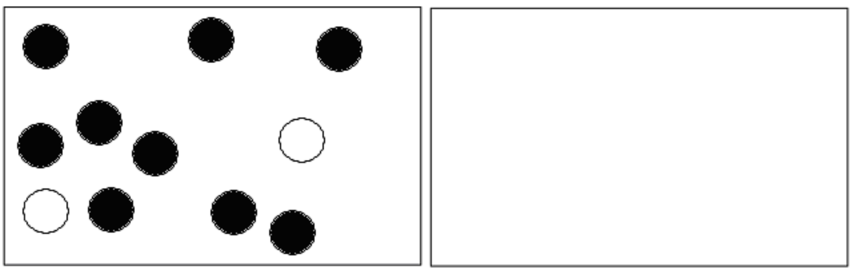


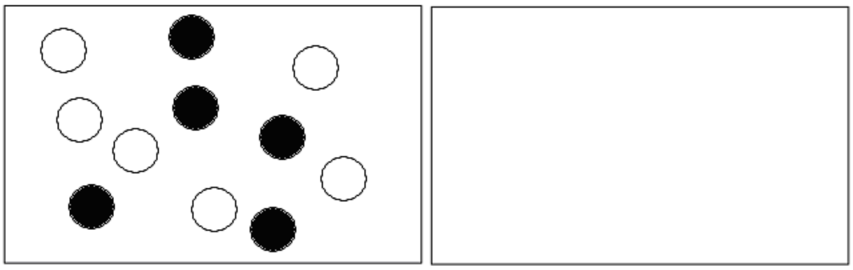


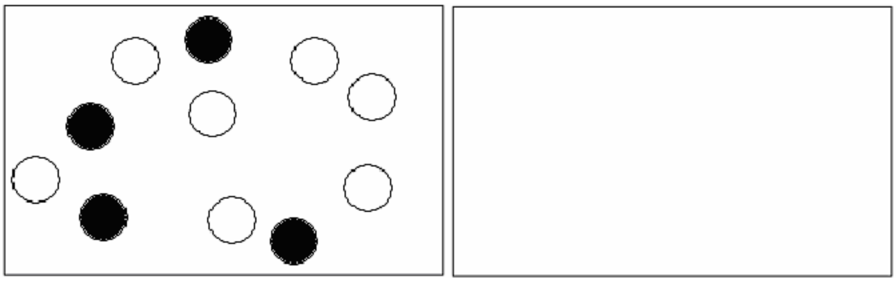


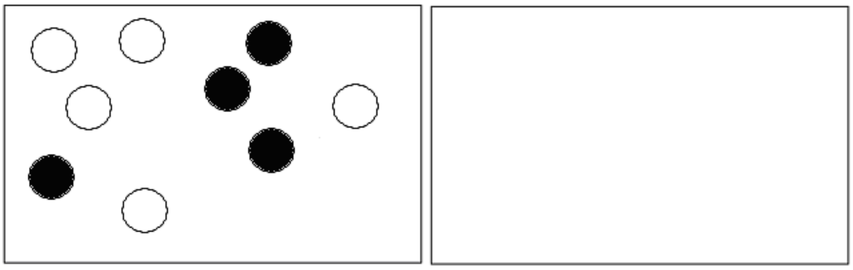


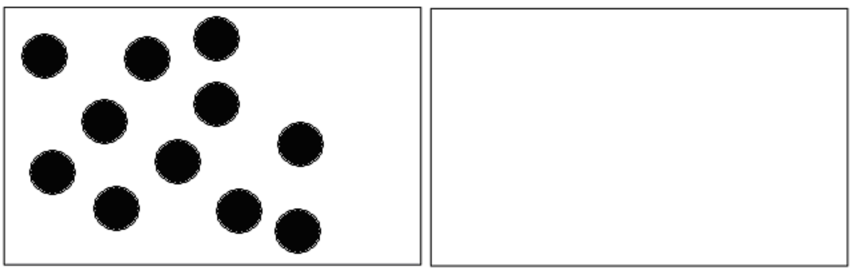


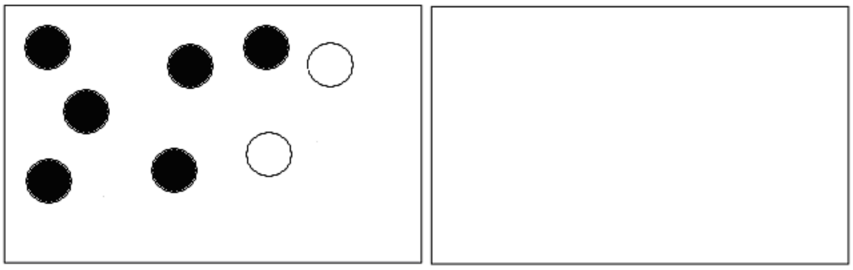


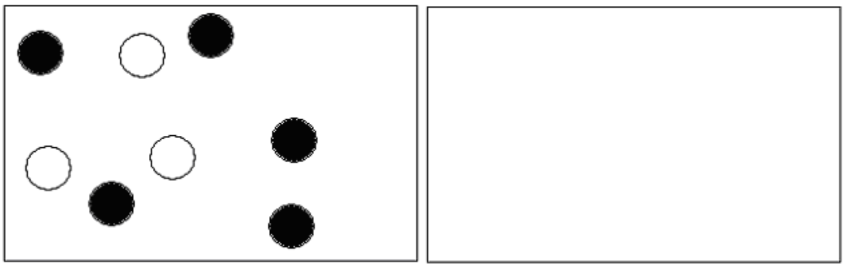


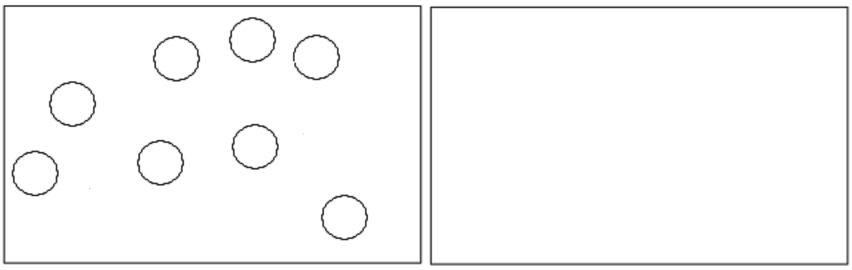


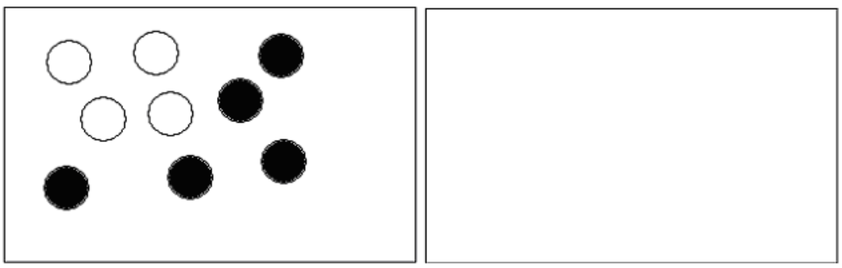


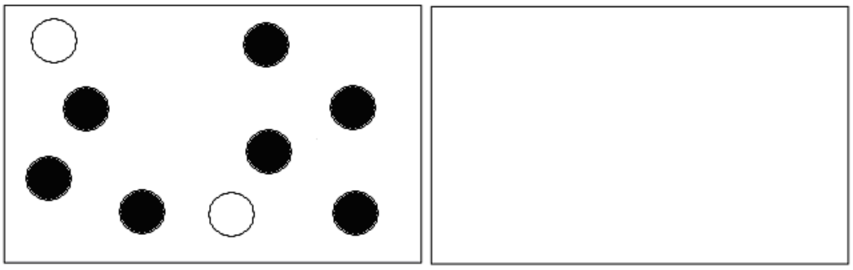


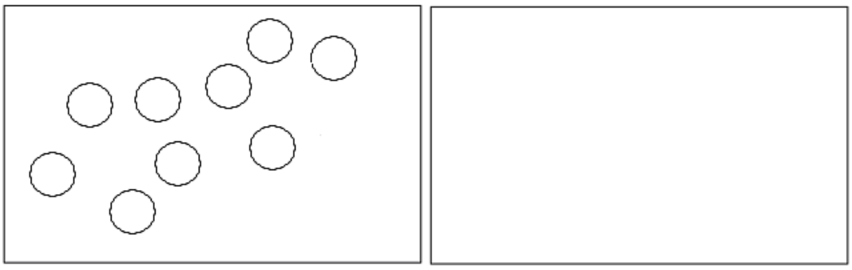


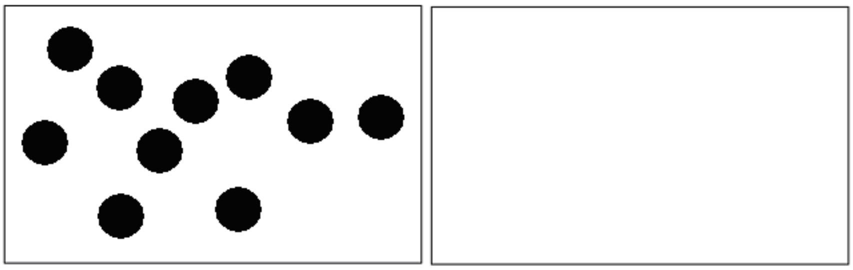


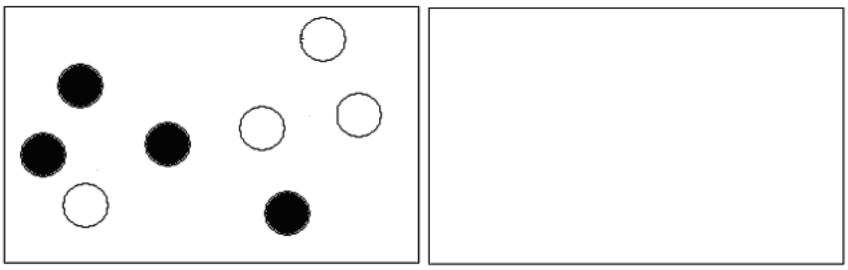












Direct inferences - empty terms

Age:

Gender:

Field of studies/level:

Tutorial:

Below you can read 12 sentences. You can infer some other sentences from them. Under each sentence write down all the sentences that in your opinion follow from the given one. Your sentences should begin with words **"all"**, **"most"**, **"no"**, **"some"** and can be both **positive** or **negative**. There can be **more than one** such inference for each given sentence. If you think that you cannot infer anything, draw a **line**.

Tasks:

1. Most mermoglins are pink.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between mermoglins and pink colour (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

2. All szaruchy are vicious.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between szaruchy and being vicious (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

3. Some puppishants are not friendly.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between puppishants and being friendly (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

4. No mroczniaki wear scarfs.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between mroczniaki and scarfs (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

5. All mermoglins are quadrupeds.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between mermoglins and quadrupeds (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

6. No present queen of Poland is a widow.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between present queens of Poland and widows (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

7. Most mgłowce have a cap.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between mgłowce and caps (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

8. Some wietrzydła have claws.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between wietrzydła and claws (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

9. Most leprocracks do not have wings.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between leprocracks and wings (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

10. Some gromblins have magical power.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between gromblins and magical power (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

11. Some traclocks are not intelligent.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between traclocks and computer being intelligent (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

12. Most Sarcotrits are not witches.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between Sarcotrits and witches (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

Direct Inferences – non-empty terms

Age:

Sex:

Field of studies/level:

Tutorial:

Below you can read 12 sentences. You can infer some other sentences from them. Under each sentence write down all the sentences that in your opinion follow from the given one. Your sentences should begin with words **"all"**, **"most"**, **"no"**, **"some"** and can be both **positive** or **negative**. There can be **more than one** such inference for each given sentence. If you think that you cannot infer anything, draw a **line**

Some of the given sentences may seem to you absurd or false. Do not worry about it.

Tasks:

1. Most of philosophy students wear berets.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between philosophy students and berets (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

2. All internists have hats.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between internists and hats (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

3. Some swimmers are not runners.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between swimmers and runners (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

4. No elephants are green.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between elephants and green colour (in positive or negative sentences, beginning with “most”, “some”, “all” or “no”)?

5. All lawyers like cheese.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between lawyers and cheese (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

6. No politician has a bike.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between politicians and bikes (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

7. Most Warsaw citizens have umbrellas.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between Warsaw citizens and umbrellas (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

8. Some professors are Mozart admirers.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between professors and Mozart admirers (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

9. Most anthropologists are not fans of diving.

   What, if anything, can you say, on the basis of the above sentences, about mutual relations between anthropologists and fans of diving (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

10. Some cellists are violinists.

    What, if anything, can you say, on the basis of the above sentences, about mutual relations between cellists and violinists (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

11. Some mathematicians are not computer scientists.

    What, if anything, can you say, on the basis of the above sentences, about mutual relations between mathematicians and computer scientists (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

12. Most students in Amsterdam are not volleyball players.

    What, if anything, can you say, on the basis of the above sentences, about mutual relations between students in Amsterdam and volleyball players (in positive or negative sentences, beginning with "most", "some", "all" or "no")?

Two-premise test – non-empty terms

Age:

Gender:

Field of study (level):

Tutorial:

Below you can read 20 examples of different pairs of sentences. From these pairs of sentences you can infer another sentences. For instance if you know that:

All smokers are mathematicians.

All Lithuanians are smokers.

You can tell that:

All Lithuanians are mathematicians.

If you think you can tell something else about mutual relations between Lithuanians and mathematicians, you can write it down here...

Read carefully the below examples. Under each pair of sentences write down all those sentences (concerning the given relation) which you think can be inferred from a given pair. Your sentences should begin with the words **"all"**, **"most"**, **"some"**, or **"no"** and can be **positive** or **negative**.

Some of the sentences may seem to you false or absurd in the light of your knowledge about the world, e.g. "All dentists like mandarins". Do not worry about it. Try to reason only on the basis of the given sentences and not on the basis of your knowledge.

Remember that it may be possible to infer **more than one conclusion** from a given pair of sentences. If, in your opinion, you cannot infer anything, draw a **line**!

Tasks:

1. You know that:

   No alpinist is a fan of sailing.

   Most logicians are alpinists.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between logicians and fans of sailing?

2. You know that:

   No scientific book has a green cover.

   Most books are not scientific books.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between books and green covers?

3. You know that:

   No firemen are doctors.

   Most Alaska inhabitants are not doctors.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between Alaska inhabitants and firemen?

4. You know that:

   All photographers are philatelists.

   Most Warsaw University employees are not philatelists.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between Warsaw University employees and photographers?

5. You know that:

   Most German philosophers are metaphysical idealists.

   Most Germans philosophers are not logicians.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between metaphysical idealists and logicians?

6. You know that:

   No Georgians have green bikes.

   Most philologists have green bikes.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between Georgians and philologists?

7. You know that:

   All dentists have three children.

   No Madagascarian has three children.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between Madagascarian and dentists?

8. You know that:

   Most lawyers have orange cars.

   Most lawyers are pragmatists.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between pragmatists and orange cars?

9. You know that:

   All musicians have notebooks

   All Latvians are musicians.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between Latvians and notebooks?

10. You know that:

    No student is a fan of waking up early.

    Most students are not singers.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between singers and fans of waking up early?

11. You know that:

    No doctor is a writer.

    All general practitioners are doctors.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between general practitioners and writers?

12. You know that:

    All algebraists are mathematicians.

    Most algebraists are not French.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between French and mathematicians?

13. You know that:

    All feminists have straw hats.

    Most women are not feminists.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between women and straw hats?

14. You know that:

    All chemists are flutists.

    Most florists are not flutists.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between florists and chemists?

15. You know that:

    Most politicians have cookery-books

    All politicians are corrupt

    What (if anything) can you tell on the basis of these sentences about the mutual relations between being corrupt and having a cookery-book?

16. You know that:

    No Slovak is a plumber.

    Most Slovaks are ecologists.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between ecologists and plumbers?

17. You know that:

    No biologist is an expert on Plato.

    Most experts on Plato are not fencers.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between biologists and fencers?

18. You know that:

    Most Englishmen are not dentists.

    All dentists are fans of mandarins.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between fans of mandarins and Englishmen?

19. You know that:

    All Chinese are ethnographers.

    Most Asians are Chinese.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between Asians and ethnographers?

20. You know that:

    No Lithuanian is Australian.

    Most Australians are owners of kangaroos.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between owners of kangaroos and Lithuanians?

Two-premise test – empty terms

Age:

Gender:

Field of study (level):

Tutorial:

Below you can read 20 examples of different pairs of sentences. From these pairs of sentences you can infer another sentences. For instance if you know that:

All dwarves have beards.

All gnomes are dwarves.

You can tell that:

All gnomes have beards.

If you think you can tell something else about mutual relations between gnomes and beards, you can write it down here...

Read carefully the below examples. Under each pair of sentences write down all those sentences (concerning the given relation) which you think can be inferred from a given pair. Your sentences should begin with the words **"all"**, **"most"**, **"some"**, or **"no"** and can be **positive** or **negative**.

Remember that it may be possible to infer **more than one conclusion** from a given pair of sentences. If, in your opinion, you cannot infer anything, draw a **line**!

Tasks:

1. You know that:

   No creatures living in underground holes are friendly.

   Most pupiszatka are creatures living in underground holes.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between pupiszatka and being friendly?

2. You know that:

   No dwarf has a blue cap.

   Most leprokraki are not dwarves.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between leprokraki and blue caps?

3. You know that:

   No wizard is defenceless.

   Most grombliny are not defenceless.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between grombliny and wizards?

4. You know that:

   All wietrzydła are dangerous.

   Mostmgłowce are dangerous.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between mgłowce and wietrzydła?

5. You know that:

   Most trakloki are beings with magical power.

   Most trakloki are not good archers.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between beings with magical power and archers?

6. You know that:

   No flying elephants know Polish.

   Most pteroklaki know Polish.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between pteroklaki and flying elephants?

7. You know that:

   All klumpiaki have violet scarfs.

   No szaruchy have violet scarfs.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between szaruchy and klumpiaki?

8. You know that:

   Most beings living on Venus planet are not Venusians.

   Most beings living on Venus planet are intelligent.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between Venusians and being intelligent?

9. You know that:

   All beings living on Mars have blue skin.

   All Martians are beings living on Mars.

   What (if anything) can you tell on the basis of these sentences about the mutual relations between Martians and blue skin?

10. You know that:

    No witch has pink gloves.

    Most witches are not Zarkotrytki.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between Zarkotrytki and pink gloves?

11. You know that:

    No creatures eating sand are Liliputs.

    All sandmen are creatures eating sand.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between sandmen and Liliputs?

12. You know that:

    All squares are quadrangles.

    Most squares are not square circles.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between square circles and quadrangles?

13. You know that:

    All singing cats bring bad luck.

    Most mroczniaki are not singing cats.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between mroczniaki and bringing bad luck?

14. You know that:

    All talking frogs are enchanted queens.

    Most living queens of Poland are not enchanted queens.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between living queens of Poland and talking frogs?

15. You know that:

    Most leprokraki have mischievous beings.

    All leprokraki have a goblet of gold.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between mischievous beings and having a goblet of gold?

16. You know that:

    No mermogliny have silver horseshoes.

    Most mermogliny are flying horses.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between flying horses and having silver horseshoes?

17. You know that:

    No archer is a gromblin.

    Most grombliny are not goblins.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between goblins and archers?

18. You know that:

    Most beings living on Never-planet are not swampmen.

    All swampmen have swamp hildren.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between living on Never-planet and having swampchildren?

19. You know that:

    All flying quadruped have sharp teeth.

    Most mermogliny are flying quadrupeds.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between mermogliny and having sharp teeth?

20. You know that:

    No elephant is a flying quadruped.

    Most flying quadrupeds are pteroklaki.

    What (if anything) can you tell on the basis of these sentences about the mutual relations between pteroklaki and elephants?

Appendix C

Tests from the second experiment

Test 1 a

Gender:

Age:

Field of study (level):

Instructions:

Below you can read 60 question. Read them carefully and based on your intuition mark YES or NO.

1. You know that *some buzaki are green.*

   Can you say, on the basis of the above sentence, that *some buzaki are not green*?

   YES / NO

2. You know that *most wyszczyki are two-colored.*

   Can you say, on the basis of the above sentence, that *some wyszczyki are two-colored*?

   YES / NO

3. You know that *most mgłowce do not have caps.*

   Can you say, on the basis of the above sentence, that *some mgłowce do not have caps*?

   YES / NO

4. You know that *some buzaki are green.*

   Can you say, on the basis of the above sentence, that *most buzaki are not green*?

   YES / NO

5. You know that *most mermogliny are pink.*

   Can you say, on the basis of the above sentence, that *some mermogliny are pink*?

   YES / NO

6. You know that *some mermogliny are pink.*

   Can you say, on the basis of the above sentence, that *most mermogliny are not pink*?

   YES / NO

7. You know that *most buzaki are not green.*

   Can you say, on the basis of the above sentence, that *some buzaki are not green*?

   YES / NO

8. You know that *most wyszczyki are not two-colored.*

   Can you say, on the basis of the above sentence, that *most wyszczyki are two-colored*?

   YES / NO

9. You know that *some leprokraki do not like cheese.*

   Can you say, on the basis of the above sentence, that *most leprokraki like cheese*?

   YES / NO

10. You know that *some zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki do not have long ears*?

    YES / NO

11. You know that *some mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *most mgłowce do not have caps*?

    YES / NO

12. You know that *most klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *some klawuchy do not have red tails*?

    YES / NO

13. You know that *most leprokraki do not like cheese.*

    Can you say, on the basis of the above sentence, that *most leprokraki like cheese*?

    YES / NO

14. You know that *some mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *most mermogliny are pink*?

    YES / NO

15. You know that *most leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki like cheese*?

    YES / NO

16. You know that *most mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *some mgłowce have caps*?

    YES / NO

17. You know that *most klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *some klawuchy have red tails*?

    YES / NO

18. You know that *most mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *most mermogliny are pink*?

    YES / NO

19. You know that *some grombliny do not have claws.*

    Can you say, on the basis of the above sentence, that *most grombliny have claws*?

    YES / NO

20. You know that *some wyszczyki are not two-colored.*

    Can you say, on the basis of the above sentence, that *most wyszczyki are two-colored*?

    YES / NO

21. You know that *most grombliny have claws.*

    Can you say, on the basis of the above sentence, that *most grombliny do not have claws*?

    YES / NO

22. You know that *most zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki have long ears*?

    YES / NO

23. You know that *most grombliny do not have claws.*

    Can you say, on the basis of the above sentence, that *some grombliny do not have claws*?

    YES / NO

24. You know that *most mroczniaki do not bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki do not bring bad luck*?

    YES / NO

25. You know that *most zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *some zarkotki do not have long ears*?

    YES / NO

26. You know that *some mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *some mgłowce do not have caps*?

    YES / NO

27. You know that *most grombliny have claws.*

    Can you say, on the basis of the above sentence, that *some grombliny have claws*?

    YES / NO

28. You know that *most leprokraków do not like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki do not like cheese*?

    YES / NO

29. You know that *some klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *most klawuchy have rad tails*?

    YES / NO

30. You know that *some trakloki are not intelligent.*

    Can you say, on the basis of the above sentence, that *most trakloki are intelligent*?

    YES / NO

31. You know that *some grombliny have claws.*

    Can you say, on the basis of the above sentence, that *most grombliny do not have claws*?

    YES / NO

32. You know that *most buzaki are green.*

    Can you say, on the basis of the above sentence, that *most buzaków are not green*?

    YES / NO

33. You know that *most wyszczyki are not two-colored.*

    Can you say, on the basis of the above sentence, that *some wyszczyki are not two-colored*?

    YES / NO

34. You know that *some mroczaniki bring bad luck.*

    Can you say, on the basis of the above sentence, that *most mroczniaki do not bring bad luck*?

    YES / NO

35. You know that *most trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are intelligent*?

    YES / NO

36. You know that *most mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *most mgłowce do not have caps*?

    YES / NO

37. You know that *some klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *most klawuchy do not have red tails*?

    YES / NO

38. You know that *some zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki have long ears*?

    YES / NO

39. You know that *some grombliny have claws.*

    Can you say, on the basis of the above sentence, that *some grombliny do not have claws*?

    YES / NO

40. You know that *most mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki bring bad luck*?

    YES / NO

41. You know that *most klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *most klawuchy have red tails*?

    YES / NO

42. You know that *some leprokraki do not like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki like cheese*?

    YES / NO

43. You know that *most zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *some zarkotki have long ears*?

    YES / NO

44. You know that *some wyszczyki are not two-colored.*

    Can you say, on the basis of the above sentence, that *some wyszczyki are two-colored*?

    YES / NO

45. You know that *some klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *some klawuchy have red tails*?

    YES / NO

46. You know that *some mgłowce do not have caps.*

    Can you say, on the basis of the above sentence, that *most mgłowce have caps.*

    YES / NO

47. You know that *some leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *most leprokraki do not like cheese*?

    YES / NO

48. You know that *most mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *most mroczniaki do not bring bad luck*?

    YES / NO

49. You know that *most trakloki are not intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are not intelligent*?

    YES / NO

50. You know that *some mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki do not bring bad luck*?

    YES / NO

51. You know that *some wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *most wyszczyki are not two-colored*?

    YES / NO

52. You know that *most mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *some mermogliny are not pink*?

    YES / NO

53. You know that *some zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *some zarkotki have long ears*?

    YES / NO

54. You know that *most trakloki are intellinegt.*

    Can you say, on the basis of the above sentence, that *most trakloki are not intelligent*?

    YES / NO

55. You know that *some buzaki are not green.*

    Can you say, on the basis of the above sentence, that *most buzaki are green*?

    YES / NO

56. You know that *some trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *most trakloki are not intelligent*?

    YES / NO

57. You know that *most buzaki are green.*

    Can you say, on the basis of the above sentence, that *some buzaki are green*?

    YES / NO

58. You know that *some mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *some mermogliny are pink*?

    YES / NO

59. You know that *some mroczniaki do not brink bad luck.*

    Can you say, on the basis of the above sentence, that *most mroczniaki bring bad luck*?

    YES / NO

60. You know that *some trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are not intelligent*?

    YES / NO

Test 1 b

Gender:

Age:

Field of study (level):

Instructions:

Below you can read 60 question. Read them carefully and based on your intuition mark YES or NO.

1. You know that *all mgłowce have caps.*

   Can you say, on the basis of the above sentence, that *no mgłowce have caps*?

   YES / NO

2. You know that *all grombliny have claws.*

   Can you say, on the basis of the above sentence, that *no grombliny have claws*?

   YES / NO

3. You know that *all mermogliny are pink.*

   Can you say, on the basis of the above sentence, that *some mermogliny are pink*?

   YES / NO

4. You know that *some mgłowce do not have caps.*

   Can you say, on the basis of the above sentence, that *some mgłowce have caps*?

   YES / NO

5. You know that *some wyszczyki are not two-colored.*

   Can you say, on the basis of the above sentence, that *no wyszczyki are two-colored*?

   YES / NO

6. You know that *some grombliny have claws.*

   Can you say, on the basis of the above sentence, that *some grombliny do not have claws*?

   YES / NO

7. You know that *some grombliny do not have claws.*

   Can you say, on the basis of the above sentence, that *some grombliny have claws*?

   YES / NO

8. You know that *all mroczniaki bring bad luck.*

   Can you say, on the basis of the above sentence, that *no mroczniaki bring bad luck*?

   YES / NO

9. You know that *no klawuchy have red tails.*

   Can you say, on the basis of the above sentence, that *some klawuchy do not have red tails*?

   YES / NO

10. You know that *some wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *some wyszczyki are not two-colored*?

    YES / NO

11. You know that *some mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *some mgłowce do not have caps*?

    YES / NO

12. You know that *no zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *some zarkotki do not have long ears*?

    YES / NO

13. You know that *no leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *all leprokraki like cheese*?

    YES / NO

14. You know that *some trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *all trakloki are intelligent*?

    YES / NO

15. You know that *no mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *all mermogliny are pink*?

    YES / NO

16. You know that *some buzaki are green.*

    Can you say, on the basis of the above sentence, that *some buzaki are not green*?

    YES / NO

17. You know that *some mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki do not bring bad luck*?

    YES / NO

18. You know that *some mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *no mermogliny are pink*?

    YES / NO

19. You know that *no mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *some mgłowce do not have caps*?

    YES / NO

20. You know that *no wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *all wyszczyki are two-colored*?

    YES / NO

21. You know that *some buzaki are green.*

    Can you say, on the basis of the above sentence, that *all buzaki are green*?

    YES / NO

22. You know that *no zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *all zarkotki have long ears*?

    YES / NO

23. You know that *no wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *some wyszczyki are not two-colored*?

    YES / NO

24. You know that *all buzaki are green.*

    Can you say, on the basis of the above sentence, that *no buzaki are green*?

    YES / NO

25. You know that *some wyszczyki are not two-colored.*

    Can you say, on the basis of the above sentence, that *some wyszczyki are two-colored*?

    YES / NO

26. You know that *some mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *some mermogliny are pink*?

    YES / NO

27. You know that *all grombliny have claws.*

    Can you say, on the basis of the above sentence, that *some grombliny have claws*?

    YES / NO

28. You know that *some trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are not intelligent*?

    YES / NO

29. You know that *some leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki do not like cheese*?

    YES / NO

30. You know that *some zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *some zarkotki do not have long ears*?

    YES / NO

31. You know that *all wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *some wyszczyki are two-colored*?

    YES / NO

32. You know that *some klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *some klawuchy have red tails*?

    YES / NO

33. You know that *some grombliny have claws.*

    Can you say, on the basis of the above sentence, that *all grombliny have claws*?

    YES / NO

34. You know that *all leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki like cheese*?

    YES / NO

35. You know that *all klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *some klawuchy have red tails*?

    YES / NO

36. You know that *some mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *some mermogliny are not pink*?

    YES / NO

37. You know that *some mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *all mgłowce have caps*?

    YES / NO

38. You know that *some klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *no klawuchy have red tails*?

    YES / NO

39. You know that *all mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *some mgłowce have caps*?

    YES / NO

40. You know that *all mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki bring bad luck*?

    YES / NO

41. You know that *some leprokraki do not like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki like cheese*?

    YES / NO

42. You know that *all trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are intelligent*?

    YES / NO

43. You know that *some leprokraki do not like cheese.*

    Can you say, on the basis of the above sentence, that *no leprokraki like cheese*?

    YES / NO

44. You know that *some mroczniaki do not bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki bring bad luck*?

    YES / NO

45. You know that *no mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki do not bring bad luck*?

    YES / NO

46. You know that *no mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *some mermogliny are not pink*?

    YES / NO

47. You know that *no trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are not intelligent*?

    YES / NO

48. You know that *all trakloki are intelligent*

    Can you say, on the basis of the above sentence, that *no trakloki are intelligent*?

    YES / NO

49. You know that *some mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *all mroczniaki bring bad luck*?

    YES / NO

50. You know that *all buzaki are green.*

    Can you say, on the basis of the above sentence, that *some buzaki are green*?

    YES / NO

51. You know that *some buzaki are not green.*

    Can you say, on the basis of the above sentence, that *some buzaki are green*?

    YES / NO

52. You know that *all zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *some zarkotki have long ears*?

    YES / NO

53. You know that *no leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki do not like cheese*?

    YES / NO

54. You know that *no grombliny have claws.*

    Can you say, on the basis of the above sentence, that *some grombliny do not have claws*?

    YES / NO

55. You know that *some zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *some zarkotki have long ears*?

    YES / NO

56. You know that *no klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *all klawuchy have red tails*?

    YES / NO

57. You know that *some zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *no zarkotki have long ears*?

    YES / NO

58. You know that *some trakloki are not intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are intelligent*?

    YES / NO

59. You know that *no buzaki are green.*

    Can you say, on the basis of the above sentence, that *some buzaki are not green*?

    YES / NO

60. You know that *some klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *some klawuchy do not have red tails*?

    YES / NO

Test 2 a

Gender:

Age:

Field of study (level):

Instructions:

Below you can read 60 question. Read them carefully and basing on your intuition mark YES or NO.

1. You know that *most buzaki are green.*

   Can you say, on the basis of the above sentence, that *all buzaki are green*?

   YES / NO

2. You know that *most zarkotki do not have long ears.*

   Can you say, on the basis of the above sentence, that *some zarkotki have long ears*?

   YES / NO

3. You know that *most mermogliny are not pink.*

   Can you say, on the basis of the above sentence, that *some mermogliny are pink*?

   YES / NO

4. You know that *most wyszczyki are not two-colored.*

   Can you say, on the basis of the above sentence, that *some wyszczyki are two-colored*?

   YES / NO

5. You know that *most leprokraki do not like cheese.*

   Can you say, on the basis of the above sentence, that *no leprokraki like cheese*?

   YES / NO

6. You know that *no buzaki are green.*

   Can you say, on the basis of the above sentence, that *most buzaki are not green*?

   YES / NO

7. You know that *no grombliny have claws.*

   Can you say, on the basis of the above sentence, that *most grombliny do not have claws*?

   YES / NO

8. You know that *all mroczniaki bring bad luck.*

   Can you say, on the basis of the above sentence, that *most mroczniaki bring bad luck*?

   YES / NO

9. You know that *no klawuchy have red tails.*

   Can you say, on the basis of the above sentence, that *most klawuchy have red tails*?

   YES / NO

10. You know that *no leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *most leprokraki like cheese*?

    YES / NO

11. You know that *all grombliny have claws.*

    Can you say, on the basis of the above sentence, that *most grombliny have claws*?

    YES / NO

12. You know that *all buzaki are green.*

    Can you say, on the basis of the above sentence, that *most buzaki are not green*?

    YES / NO

13. You know that *all zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki have long ears*?

    YES / NO

14. You know that *most trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *all trakloki are intelligent*?

    YES / NO

15. You know that *all mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *most mgłowce have caps*?

    YES / NO

16. You know that *most klawuchów have red tails.*

    Can you say, on the basis of the above sentence, that *some klawuchy do not have red tails*?

    YES / NO

17. You know that *most mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *all mroczniaki bring bad luck*?

    YES / NO

18. You know that *all klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *most klawuchy have red tails*?

    YES / NO

19. You know that *no wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *most wyszczyki are two-colored*?

    YES / NO

20. You know that *most mgłowce do not have caps.*

    Can you say, on the basis of the above sentence, that *some mgłowce have caps*?

    YES / NO

21. You know that *most mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *some mermogliny are not pink*?

    YES / NO

22. You know that *most zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *no zarkotki have long ears*?

    YES / NO

23. You know that *all leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *most leprokraki like cheese*?

    YES / NO

24. You know that *all buzaki are green.*

    Can you say, on the basis of the above sentence, that *most buzaki are green*?

    YES / NO

25. You know that *most grombliny have claws.*

    Can you say, on the basis of the above sentence, that *some grombliny do not have claws*?

    YES / NO

26. You know that *most traklokÓw are not intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are intelligent*?

    YES / NO

27. You know that *most mroczniaki do not bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki bring bad luck*?

    YES / NO

28. You know that *most leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki do not like cheese*?

    YES / NO

29. You know that *no mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *most mgłowce do not have caps*?

    YES / NO

30. You know that *no wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *most wyszczyki are not two-colored*?

    YES / NO

31. You know that *all trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *most traklokо́w are not intelligent*?

    YES / NO

32. You know that *all mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *most mroczniaki do not bring bad luck*?

    YES / NO

33. You know that *most mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *all mgłowce have caps*?

    YES / NO

34. You know that *no klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *most klawuchy do not have red tails*?

    YES / NO

35. You know that *no leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *most leprokraki do not like cheese*?

    YES / NO

36. You know that *no mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *most mermogliny are not pink*?

    YES / NO

37. You know that *no trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *most trakloki are not intelligent*?

    YES / NO

38. You know that *most wyszczyki are not two-colored.*

    Can you say, on the basis of the above sentence, that *no wyszczyki are two-colored*?

    YES / NO

39. You know that *most leprokraki do not like cheese.*

    Can you say, on the basis of the above sentence, that *some leprokraki like cheese*?

    YES / NO

40. You know that *all mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *most mermogliny are pink*?

    YES / NO

41. You know that *most klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *no klawuchy have red tails*?

    YES / NO

42. You know that *all grombliny have claws.*

    Can you say, on the basis of the above sentence, that *most grombliny do not have claws*?

    YES / NO

43. You know that *no mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *most mroczniaki do not bring bad luck*?

    YES / NO

44. You know that *no mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *most mermogliny are pink*?

    YES / NO

45. You know that *no zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki have long ears*?

    YES / NO

46. You know that *most grombliny do not have claws.*

    Can you say, on the basis of the above sentence, that *some grombliny have claws*?

    YES / NO

47. You know that *most mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *some mgłowce do not have caps*?

    YES / NO

48. You know that *most klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *some klawuchy have red tails*?

    YES / NO

49. You know that *most grombliny have claws.*

    Can you say, on the basis of the above sentence, that *all grombliny have claws*?

    YES / NO

50. You know that *no zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki do not have long ears*?

    YES / NO

51. You know that *most mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *some mroczniaki do not bring bad luck*?

    YES / NO

52. You know that *most zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *some zarkotki do not have long ears*?

    YES / NO

53. You know that *all wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *most wyszczyki are two-colored*?

    YES / NO

54. You know that *most wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *some wyszczyki are not two-colored*?

    YES / NO

55. You know that *most mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *no mermogliny are pink*?

    YES / NO

56. You know that *all trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *most trakloki are intelligent*?

    YES / NO

57. You know that *all mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *most mgłowce do not ave caps*?

    YES / NO

58. You know that *most buzaki are green.*

    Can you say, on the basis of the above sentence, that *some buzaki are not green*?

    YES / NO

59. You know that *most trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *some trakloki are not intelligent*?

    YES / NO

60. You know that *most buzaki are not green.*

    Can you say, on the basis of the above sentence, that *some buzaki are green*?

    YES / NO

Test 2 b

Gender:

Age:

Field of study (level):

Instructions:

Below you can read 60 question. Read them carefully and basing on your intuition mark YES or NO.

1. You know that *no mermogliny are pink.*

   Can you say, on the basis of the above sentence, that *most mermogliny are pink*?

   YES / NO

2. You know that *all buzaki are green.*

   Can you say, on the basis of the above sentence, that *most buzaki are green*?

   YES / NO

3. You know that *all mroczniaki bring bad luck.*

   Can you say, on the basis of the above sentence, that *most mroczniaki do not bring bad luck*?

   YES / NO

4. You know that *most grombliny have claws.*

   Can you say, on the basis of the above sentence, that *not all grombliny have claws*?

   YES / NO

5. You know that *no mgłowce do not have caps.*

   Can you say, on the basis of the above sentence, that *most mgłowce do not have caps*?

   YES / NO

6. You know that *most zarkotki have long ears.*

   Can you say, on the basis of the above sentence, that *not all zarkotki have long ears*?

   YES / NO

7. You know that *all mermogliny are pink.*

   Can you say, on the basis of the above sentence, that *most mermogliny are pink*?

   YES / NO

8. You know that *most mroczniaki bring bad luck.*

   Can you say, on the basis of the above sentence, that *not all mroczniaki bring bad luck*?

   YES / NO

9. You know that *all mgłowce have caps.*

   Can you say, on the basis of the above sentence, that *most mgłowce do not have caps*?

   YES / NO

10. You know that *most trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *all trakloki are intelligent*?

    YES / NO

11. You know that *no wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *most wyszczyki are not two-colored*?

    YES / NO

12. You know that *most grombliny do not have claws.*

    Can you say, on the basis of the above sentence, that *it is not the case that no grombliny have claws*?

    YES / NO

13. You know that *all grombliny have claws.*

    Can you say, on the basis of the above sentence, that *most grombliny have claws*?

    YES / NO

14. You know that *most klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *it is not the case that no klawuchy have red tails*?

    YES / NO

15. You know that *most wyszczyki are not two-colored.*

    Can you say, on the basis of the above sentence, that *no wyszczyki are two-colored*?

    YES / NO

16. You know that *most mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *it is not the case that no mermogliny are pink*?

    YES / NO

17. You know that *most klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *not all klawuchy have red tails*?

    YES / NO

18. You know that *most mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *all mgłowce have caps*?

    YES / NO

19. You know that *most buzaki are green.*

    Can you say, on the basis of the above sentence, that *all buzaki are green*?

    YES / NO

20. You know that *no leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *most leprokraki do not like cheese*?

    YES / NO

21. You know that *most mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *not all mgłowce have caps*?

    YES / NO

22. You know that *no wyszczyki not are two-colored.*

    Can you say, on the basis of the above sentence, that *most wyszczyków are two-colored*?

    YES / NO

23. You know that *all trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *most traklokόw not are intelligent*?

    YES / NO

24. You know that *no leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *most leprokraków like cheese*?

    YES / NO

25. You know that *most mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *all mroczniaki bring bad luck*?

    YES / NO

26. You know that *most zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *no zarkotki have long ears*?

    YES / NO

27. You know that *most wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *not all wyszczyki are two-colored*?

    YES / NO

28. You know that *all mgłowce have caps.*

    Can you say, on the basis of the above sentence, that *most mgłowce have caps*?

    YES / NO

29. You know that *most mermogliny are not pink.*

    Can you say, on the basis of the above sentence, that *no mermogliny are pink*?

    YES / NO

30. You know that *most buzaki are not green.*

    Can you say, on the basis of the above sentence, that *it is not the case that no buzaki are green*?

    YES / NO

31. You know that *all buzaki are green.*

    Can you say, on the basis of the above sentence, that *most buzaki are not green*?

    YES / NO

32. You know that *most zarkotki do not have long ears.*

    Can you say, on the basis of the above sentence, that *it is not the case that no zarkotki have long ears*?

    YES / NO

33. You know that *most mroczniaki do not bring bad luck.*

    Can you say, on the basis of the above sentence, that *it is not the case that no mroczniaki bring bad luck*?

    YES / NO

34. You know that *no buzaki are green.*

    Can you say, on the basis of the above sentence, that *most buzaki are not green*?

    YES / NO

35. You know that *all trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *most trakloki are intelligent*?

    YES / NO

36. You know that *most wyszczyków are not two-colored.*

    Can you say, on the basis of the above sentence, that *it is not the case that no wyszczyki are two-colored*?

    YES / NO

37. You know that *most leprokraki do not like cheese.*

    Can you say, on the basis of the above sentence, that *it is not the case that no leprokraki like cheese*?

    YES / NO

38. You know that *all grombliny have claws.*

    Can you say, on the basis of the above sentence, that *most grombliny do not have claws*?

    YES / NO

39. You know that *most buzaki are green.*

    Can you say, on the basis of the above sentence, that *not all buzaki are green*?

    YES / NO

40. You know that *no mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *most mermogliny are not pink*?

    YES / NO

41. You know that *no zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki do not have long ears*?

    YES / NO

42. You know that *all mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *most mroczniaki bring bad luck*?

    YES / NO

43. You know that *most mgłowce do not have caps.*

    Can you say, on the basis of the above sentence, that *it is not the case that no mgłowce have caps*?

    YES / NO

44. You know that *all wyszczyki are two-colored.*

    Can you say, on the basis of the above sentence, that *most wyszczyki are two-colored*?

    YES / NO

45. You know that *all zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki have long ears*?

    YES / NO

46. You know that *no zarkotki have long ears.*

    Can you say, on the basis of the above sentence, that *most zarkotki have long ears*?

    YES / NO

47. You know that *all leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *most leprokraki like cheese*?

    YES / NO

48. You know that *no trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *most trakloki are not intelligent*?

    YES / NO

49. You know that *most leprokraki like cheese.*

    Can you say, on the basis of the above sentence, that *not all leprokraki like cheese*?

    YES / NO

50. You know that *most trakloki are intelligent.*

    Can you say, on the basis of the above sentence, that *not all trakloki are intelligent*?

    YES / NO

51. You know that *no klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *most klawuchy have red tails*?

    YES / NO

52. You know that *no klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *most klawuchy do not have red tails*?

    YES / NO

53. You know that *most grombliny have claws.*

    Can you say, on the basis of the above sentence, that *all grombliny have claws*?

    YES / NO

54. You know that *no mroczniaki bring bad luck.*

    Can you say, on the basis of the above sentence, that *most mroczniaki do not bring bad luck*?

    YES / NO

55. You know that *most leprokraki do not like cheese.*

    Can you say, on the basis of the above sentence, that *no leprokraki like cheese*?

    YES / NO

56. You know that *most mermogliny are pink.*

    Can you say, on the basis of the above sentence, that *not all mermogliny are pink*?

    YES / NO

57. You know that *no grombliny have claws.*

    Can you say, on the basis of the above sentence, that *most grombliny do not have claws*?

    YES / NO

58. You know that *most klawuchy do not have red tails.*

    Can you say, on the basis of the above sentence, that *no klawuchy have red tails*?

    YES / NO

59. You know that *all klawuchy have red tails.*

    Can you say, on the basis of the above sentence, that *most klawuchów have red tails*?

    YES / NO

60. You know that *most trakloki are not intelligent.*

    Can you say, on the basis of the above sentence, that *it is not the case that no trakloki are intelligent*?

    YES / NO

Bibliography

Barwise K.J., *On branching quantifiers in English*, Journal of Philosophical Logic, Vol 8, Nr 1, 1979.

Barwise J., Cooper R. *Generalized Quantifiers in natural language*, Linguistics and Philosophy 4(1):159–219, 1981.

Bucholc P., *Kompetencja logiczna a poprawność logiczna. Analiza na przykładzie terminów pustych*, (*Logical competence and logical correctness – analysis on the basis of empty terms*) [in:] *Kognitywistyka*, Koło Filozoficzne Przy MISH, Warszawa, 2004.

Chater N., Oaksford M., *The Probability Heuristics Model of Syllogistic Reasoning*, Cognitive Psychology, 38, 191-258, 1999.

Field A., *Discovering statistics using SPSS*, Sage, 2009.

Grice H.P., *Studies in the Ways of Words*, Cambridge, MA: Harvard University Press, 1989.

Kamp H., *A theory of truth and semantic representation*, [in:] J.A.G. Groenendijk, T.M.V. Janssen, and M.B.J. Stokhof (eds.), Formal Methods in the Study of Language. Mathematical Centre Tracts 135, Amsterdam, 1981.

Kant Immanuel, *Critique of Pure Reason*, trans. Norman Kemp Smith, 2006.

Kaufman E. L., Lord, M. W., Reese T. W., & Volkmann J, *The discrimination of visual number*, American Journal of Psychology 62: 498–525, 1949.

Lambalgen M. van, Stenning K. *Human Reasoning and Cognitive Science*, 2008.

Levinson, S. C., *Pragmatics* Cambridge, England: Cambridge University, 1983.

Lindström, P., *First order predicate logic with generalized quantifiers.* Theoria, 32: 186-195, 1966.

Łukasiewicz J., *Elements of mathematical logic*, Translated from Polish by Olgierd Wojtasiewicz. New York, Macmillan 1964.

Łukasiewicz J., *Z zagadnień logiki i filozofii. Pisma Wybrane.* (*Selected Works*), J.Słupecki (red.), PWN, Warszawa, 1961.

Macnamara J., *A Border Dispute. The Place of Logic in Psychology.* Cambidge MA, The Massachusettes Institute of Technology, 1986.

Miller Barry, *Existence*, The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.), *URL* =< *http* : *//plato.stanford.edu/archives/fall*2008*/entries/existence/* >.

Mostowski M., Bucholc P., *O problemie terminów pustych* (*Empty terms*) [in:] Ratione et Studio, (ed.) Trzasicki K., Wydawnictwo Uniwersytetu w Białymstoku, Białystok, 2005.

Nęcka E., Orzechowski J., Szymura B., *Psychologia Poznawcza* (*Cognitive Psychology*), PWN, Warszawa, 2006.

Parsons T., *The Traditional Square of Opposition*, The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.), *URL* =< *http* : *//plato.stanford.edu/archives/fall*2008*/entries/square/* >.

Peters, S., Westerstahl, D. *Quantifiers in Langauge and Logic*, Oxford University Press, 2006.

Priest Graham, *Towards non-being. The Logic and Methaphisics of Intentionality*, Calderon Press, Oxford, 2005.

Reicher Maria, *Nonexistent Objects*, The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.), *URL* =< *http* : *//plato.stanford.edu/archives/fall*2008*/entries/nonexistent* − *objects/* >.

Rescher N., *Plurality quantification*, Journal of Symbolic Logic, 27, 1962.

Sternberg, R. J., *Cognitive Psychology*, Holt Rinehart & Winston, 1996.

Szymanik J., *Quantifiers in Time and Space*, ILLC Dissertation Series DS-2009-01.

Taves E. H., *Two mechanisms for the perception of visual numerousness*, Archives of Psychology 37: 1–47, 1941.

Trick L. M., & Pylyshyn Z. W, *Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision*, Psychological Review 101 (1): 80–102, 1994.

Abstract

The thesis discusses two properties of reasoning with quantifiers in natural language: scalar implicatures and existential import. The notion of scalar implicature refers in this case to inferring “not all” from both “some” and “most”, or “not most” from “some”; existential import – to the question whether the universal categorial sentence “All A’s are B” implies that there are A’s. Traditionally it is connected with the subalternation relation in the Aristotle’s Square of Opposition, namely with inferring “Some A’s are B” from “All A’s are B”.

We conducted a series of reasoning experiments to check the above properties with respect to people’s active inference-production (which we call active logical competence) as well as their passive evaluation of given inferences (passive logical competence). We compared reasoning in empty and non-empty domains. Based on our results and especially on the lack of any connection between subject’s inferences and emptiness of domains, we explain all the observed phenomena in terms of scalar properties of quantifiers and refuse to analyze inferences with universal premises and particular conclusions in terms of existential import. We propose to treat quantifiers “some” and “most” as vague with flexible denotations. We provide a formalization of vagueness of quantifiers in terms of fuzzy semantics, and assume that quantifiers have default denotations which can be extended in the process of reasoning. The implicature “not all” is explained by the default meaning of “most” and “some”.

To deal with the issue of interpretation of existential sentences from fictional discourse we propose to use a dynamic modification of possible worlds semantics. We treat worlds as mental constructions describing certain world states. Existential sentences presuppose a certain world or class of worlds that set up a domain of discourse.