Investigating cross-linguistic differences in the pragmatics of scalar terms[1]

Izabela Skoczeń[1] and Aleksander Smywiński-Pohl[2]

[1] Department of Legal Theory, Jagiellonian University, Cracow, Poland
[2] WIET AGH University of Technology in Cracow, Poland
izabela.skoczen@uj.edu.pl

Abstract. N. Goodman and A. Stuhlmüller made an experiment to empirically confirm a model that predicted the probability of implicature formation in contexts where speaker and hearer do not have full knowledge of the situation. However, there remained the question whether the effects were cross-linguistic or rather confined to the semantics of the English language. In our paper we present the replication of the above-described experiment in the Polish language. The data depicts differences that may point out to differences in the pragmatics of the Polish word for some, namely 'niektóre', since the scalar implicature is not cancelled in incomplete knowledge contexts.

Keywords: scalar implicature, incomplete knowledge, rational speech act model.

1 Introduction

The aim of the present paper is to answer two main questions. Firstly, we investigate whether the rational Speech Act model is cross-linguistic. Secondly, we investigate whether pragmatic behavior, namely cancelling scalar implicatures in incomplete knowledge situations occurs cross-linguistically.

1.1 Scalar implicatures

In 1975, Paul Grice, a British philosopher, introduced in his article 'Logic and Conversation' the notion of implicature (Grice, 1975). Grice noticed that through uttering sentences in context, people convey more than just the meaning of the words they use. Consider the following example:

A: Are you hungry?
B: I have had breakfast.

Imagine that this conversation takes place before noon. Although, literally taken, B's reply has not much in common with A's question, the conversation is perfectly understandable. This is because it is possible to infer from B's reply that she is not hungry.

[1] The research was funded by the National Science Centre, Poland, Preludium grant nr 2015/19/N/HS5/00029. The Authors would like to thank Leon Bergen, Daniel Lassiter, Katarzyna Kijania-Placek and Krzysztof Kasparek for very helpful comments on data analysis.

Thus, the implicature carried by B's answer will be that 'B is not hungry'. Moreover, the speaker can cancel implicatures explicitly. Consider the following example:

(I.) A: Are you hungry?

(II.) B: I have had breakfast, but a very modest one.

The second part of B's reply is an explicit cancellation of the implicature formed on the basis of the first part of the sentence, namely that 'B is not hungry'.

Scalar implicatures are special because they are based on conventional meanings attributed to words with the use of lexical scales (Horn, 2006). If you hear 'some of the students have passed the exam' you will infer that not all of them have passed.

The reason for which the 'not all; part of 'some' is pragmatic rather than semantic is that if you compare two utterances:

(III.) 'Some of the students have passed the exam', (while in fact all of them have)

(IV.) 'All of the students have passed the exam', (while in fact some of them have)

The utterer of (IV) is more likely to be accused of lying than the utterer of (III). Note that this is a contextual intuition: if you say 'some of the accused were sentenced', while in fact all of them have been sentenced, your intuition might change. For this reason, the 'not all' part of some seems contextual rather than semantic.

A lexical scale is defined as a 'totally ordered set of lexical items which vary along a single dimension'. (Bergen et al, 2015, p. 10) Placing a word in a definite position on a scale enables the speakers to attribute it a definite meaning that does not have to be consistent with the lexical meaning of the word. The most discussed example in the literature is the quantifier term "some" that can mean either 'at least some' or 'some but not all'. In ordinary conversation, 'some' is understood as 'some but not all'. Thus, we take 'at least some or some but not all' to be the lexical meaning of some. An analogous phenomenon is observed in the Polish language. The quantifier 'niektóre' (some) can mean either 'przynajmniej niektóre' (at least some) or 'tylko niektóre' (some but not all). Moreover, in an ordinary conversation 'niektóre' (some) is understood as 'tylko niektóre' (some but not all).

Consider the following example:

(V.) Some of the guests have gone.

When uttering (V.) the speaker is normally taken as implicating that not all of the guests have gone. This is because, had the speaker wanted to be understood as implicating 'at least some of the guests have gone, in fact all of them' she would have used the word 'all' rather than 'some'. The above considerations are valid in situations where the speaker and hearer have full contextual knowledge. This means that the speaker knows exactly how many of the guests have gone. She knows whether all of them have gone or only a number of them have gone. Moreover, the hearer knows that the speaker has this knowledge. Thus, the scale is the following:

<some, all>

Before proceeding further, one last remark concerning the semantics of the Polish word 'niektóre'. It is worth noting that although in example (V) the Polish word would be understood as 'some but not all' and the implicature would be cancellable, there are differences in the remaining possible uses of the English 'some'. Namely,

the English ‘some’ can be used as a singular or plural indefinite in the following sentences:

(VI.) Would you like some biscuits? Yes, I would like some.

(VII.) Yesterday, I went to the store and I bought some books.

By contrast, the Polish translation of the sentences above could not include the word ‘niektóre’. Polish users would rather employ words such as ‘jakieś’ or ‘pewne’.

1.2 The Rational Speech Act model of scalar implicatures

The scalar inference should not occur when both speaker and hearer do not have full information about the number of guests that are gone from the party. In such situation, when the speaker uses the word ‘some’ she should be understood as implicating ‘at least some’. Thus, the scalar implicature ‘some but not all of the guests are gone’ should be cancelled.

Numerals such as “two” or “twenty” are also scalar notions. This is because, lexically speaking ‘two’ could mean: ‘exactly two’, ‘at least two’ and ‘not more than two’. Analogously ‘twenty’ could mean ‘exactly twenty’, ‘at least twenty’ and ‘not more than twenty’. When the speaker has full knowledge and the hearer knows that the speaker has full knowledge then through using ‘two’ or ‘twenty’ the speaker will usually be taken as implicating ‘exactly two’ and ‘exactly twenty’. It is based on the assumption that if the speaker had meant more or less than ‘twenty’ she would have used a word that is higher on the scale (for instance ‘twenty-one’) or lower on the scale (for instance ‘nineteen’). (Horn, 2006) Thus, the scale is this:

<one, two, three, four, five, etc.>

By contrast, the implicated exactness should be cancelled when the speaker does not have full contextual knowledge. (Goodman and Stuhlmüller, 2013) Imagine a party to which three guests have been invited. The speaker knows that two of the guests have left the party, yet the speaker does not know whether the third guest has left the party. The speaker says:

(VIII.) Two of the guests have gone.

If the hearer is aware of the speaker’s partial knowledge, the hearer should cancel (or not infer) the implicature ‘exactly two’.

Noah Goodman and Andreas Stuhlmüller have created a model based on game theory and Bayesian decision theory, which aims to predict the probability of (not) cancelling the scalar implicature, in scenarios where the speaker has partial contextual knowledge and the hearer knows that the speaker has partial contextual knowledge.

The model is a ‘rational speech act model’. It views ‘language comprehension as a rational inference based on an intuitive theory of language production (...) the listener infers the world state, *s*, given the speaker’s utterance, *w*, and shared information about the speaker’s (possibly incomplete) information access, *a*. By Bayes’ rule:

$$P_{\text{listener}}(s|w,a) \propto P_{\text{speaker}}(w|s,a)P(s)$$

where P(s) captures the listener’s prior beliefs about the world state and P_{speaker} (w⁄s,a) describes the listener’s intuitive theory of how the speaker chooses words.’ (Goodman and Stuhlmüller, 2013, p. 175)

A speaker 'makes observation o about the true state of the world and selects an utterance w to convey information about the world state to a listener and does so by soft-max optimizing expected utility':

$$P_{speaker}(w|o,a) \propto \exp(\alpha E_{P(s|o,a)}[U(w;s)])$$

'The speaker's utility function, U(w;s), captures the value of saying w if the world is actually s. The expectation is taken over the speaker's belief state, P(s|o,a), because the speaker may still be uncertain about the state of the world. The parameter α controls the deviation from optimality.' (Goodman and Stuhlmüller, 2013, p. 175)

'To capture a motivation to be informative, utility must be related to the information conveyed in the utterance. More specifically, utility is related to the amount of information that a literal listener would not yet know about state s after hearing it described by utterance w (the negative surprisal):

$$U(w;s) = \ln(P_{lex}(s|w))$$

where the literal interpretation probability $P_{lex}(s|w)$ is determined by the lexicon '(Goodman and Stuhlmüller, 2013, p. 175)

The 'speaker's access a is common knowledge of speaker and listener, but the listener still does not know what observation the speaker made:

$$P_{speaker}(w|s,a) = \sum_{o} P_{speaker}(w|o,a)P(o|a,s)$$

For this reason, the hearer is performing a summing on the possible observations made by the speaker. (Goodman and Stuhlmüller, 2013, p. 175)

The model remains agnostic on the 'question under discussion', which is a restriction of the available alternatives, since this concept was introduced later (see for instance Benz et al, 2017).

This model has been empirically tested and confirmed in the English language in (Goodman and Stuhlmüller, 2013) During the experiment the speaker saw for example three letters. However, he had information concerning whether there are checks in only two of the letters. The hearer was aware of the speaker's state of knowledge. The utterance 'some of the letters have checks inside' cancelled the implicature 'not all' since hearers inferred a similar probability of two and three letters having checks. If the speaker saw two out of three letters and said 'one of the letters has a check inside', then the hearer inferred a similar probability of one and two of the letters having checks inside and a low probability of three letters having checks inside. Thus, the implicature 'not more than one' was cancelled. During the experiment six scenarios were used and the speaker had a varied access to objects and their features, while the hearer was aware of the speaker's state of knowledge.

Each scenario began with a boosting of the probability that all the objects have the salient property, for example:

> Letters to Laura almost always have checks inside. (Goodman and Stuhlmüller, 2013, p. 178)

Next, came information about the number of considered objects and a question about the prior probability distribution:

> Today Laura received three letters. How many of the 3 letters do you think have checks inside? (Goodman and Stuhlmüller, 2013, p. 178)

The participants had to bet a total of 100 dollars on each of the options, namely zero, one, two or three objects having the property. After having captured the priors came the information about the number of objects checked by the speaker:

> Laura tells you on the phone: "I have looked at 2 of the 3 letters. Some of the letters have checks inside." (Goodman and Stuhlmüller, 2013, p. 178)

Finally, the posterior probability distribution as well as a control question followed:

> Do you think Laura knows exactly how many of the 3 letters have checks inside? (Goodman and Stuhlmüller, 2013, p. 178)

The experiment confirmed a fine-grained interaction between the state of knowledge about the world and pragmatic inference. It also provided a strong argument against modular theories of mind since the data showed that the language faculty and the inference about world state were strongly connected. However, there remained the question whether the effects were cross-linguistic or rather confined to the English language.

Below we present the replication of the above-described experiment in the Polish language. We translated the same scenarios into Polish. We employed the Polish words 'niektóre' for 'some' and 'wszystkie' for 'all' as well as the numerals 'jeden, dwa, trzy' for 'one, two, three' respectively. The data proved to be a replication of Goodman and Stuhlmüller's experiment with some differences in the results. Namely, there was an interaction between the state of knowledge about the world and pragmatic inference conforming to the model predictions. However, the scalar implicature of the word some, that is 'not all' was not cancelled when the speaker had access to two out of three objects. In the paper, we investigate hypotheses that explain the differences.

2 Replication results in Polish

Methods

We carried four experiments in Polish: two that tested numerals and two that tested the quantifier 'some'. The experiments were an exact methodological replication of the original experiment. Participants were presented two control questions. In the numeral-experiment we employed 6 scenarios that appeared in a randomized order. In the quantifier experiment we also employed 6 scenarios but each participant was presented with only 3 of them in a randomized order. Each scenario was about an three objects and a potential property that these objects could have. Each participant disposed of a quantity of 100 units (zlotys) and had to bet on four possibilities: zero of the objects having the property, one of the objects having the property, two of the objects having the property and three of the objects having the property.

We tested 6 conditions: '"One" access 1' where the speaker had access to information concerning the properties of one object out of three and uttered 'one'. 'Access 2' –the speaker has access to information concerning the properties of two objects out of three and 'access 3' – the speaker has access to information concerning the properties of three objects out of three. The possible utterances were 'one', 'two' and 'three'.

The participants were recruited through an online platform ‘Research online’ and performed the experiment for a small payment. Each of the experiments was made with the participation of 50 persons – 50% men and 50% women.

Interestingly, we carried a pilot study involving 10 academics and it proved to be a neat replication of the American results. The bets that the speaker had complete knowledge in partial-access conditions were: M=18,89 SD=26,07, in complete access conditions: M=92,59 SD=21,94 compared to the original: M=42, SD=3,4 and M=92,1 SD=1,6.

The graphs present mean participant bet on each world state, varying the word the speaker used and the speaker’s perceptual access. Error bars are standard error of the mean. Statistical significance was calculated using the t-student test for dependent or independent measurement. The confidence interval (CI) was 95%. The effect size calculations were based in Cohen’s d.

Testing numerals

50 students too part in our first experiment and the results replicated (see Fig 1). The bets that the speaker had complete knowledge in partial-access conditions were: M=27,95 SD=33, in complete access conditions: M=91,79 SD=14,94 compared to the original: M=42 SD=3,4 and M=92,1 SD=1,6.

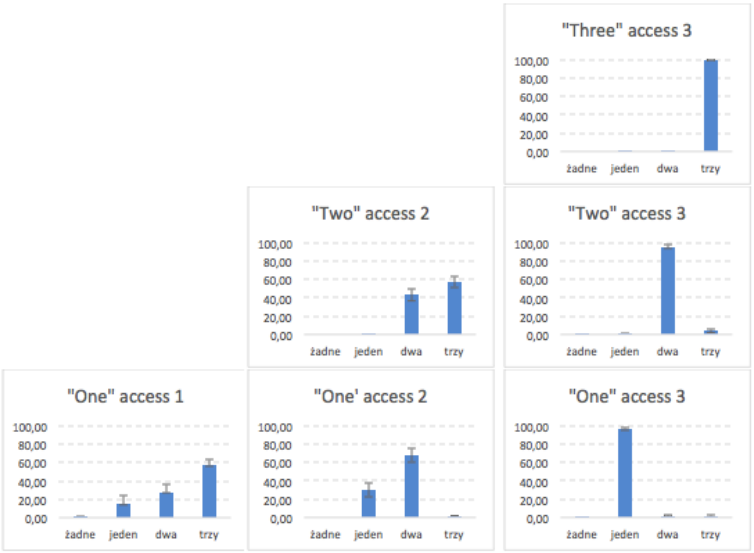

Fig 1 replication with the participation of 50 students - Mean participant bet on each world state, varying the word the speaker used and the speaker’s perceptual access. Data have been filtered to include only trials where the participant’s bet that the speaker had complete knowledge was greater than 70 in the expected direction. Error bars are standard error of the mean. The formulation ‘”One” access 1’ means that the speaker has access to information concerning the properties of one object out of three and utters ‘one’. ‘Access 2’ –the speaker has access to information concerning the properties of two objects out of three and ‘access 3’ – the speaker has access to information concerning the properties of three objects out of three.

In the experiment represented in Fig 2, the sample was modified to 20 people aged 16-29 years (10 men and 10 women), 20 people aged 30-49 (10 men and 10 women), 10 people aged over 50 (5 men and 5 women). Subjects were recruited through an online survey platform 'Research online'. Differences appeared in the 'one access one' and 'two access two' conditions. The difference in the 'two access two' conditions was crucial since in the original experiment the scalar implicature 'not more than two' was cancelled, while in the data we collected the scalar implicature persisted. In other words, the bets on 'two' and 'three' in the '"two" access 2' condition were similar in the American experiment, while in the replication in Polish the bets on 'two' were significantly higher than the bets on 'three'. The bets that the speaker had complete knowledge in partial-access conditions were: M= 48,53 SD= 37,04, in complete access conditions: M= 86,47 SD= 20,47 compared to the original: M=42 SD=3,4 and M=92,1 SD=1,6.

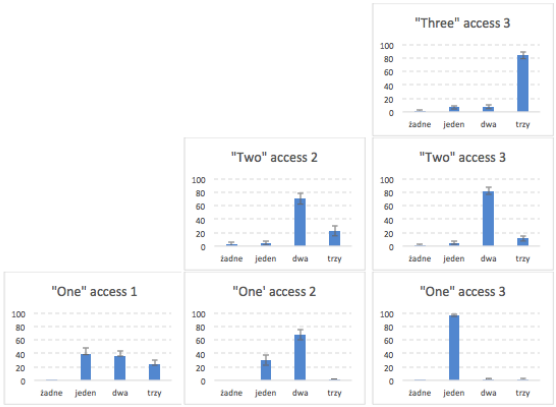

Fig 2 replication with the participation of 50 subjects from varied age groups

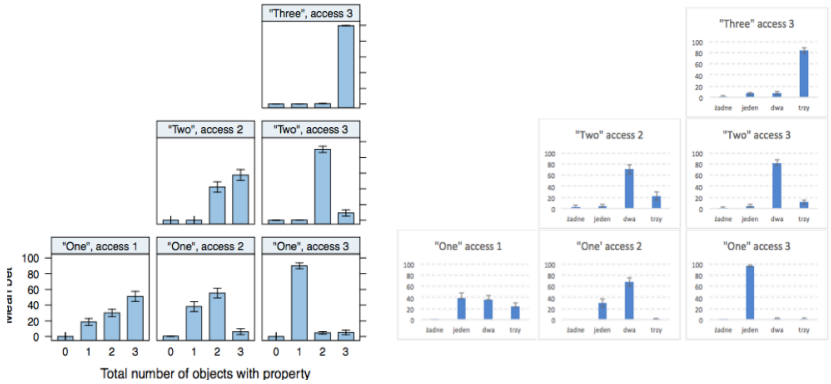

Fig 3 Differences between the original experiment (to the left, reuse license number 4343821165053, source: https://onlinelibrary.wiley.com/doi/abs/10.1111/tops.12007) and the replication (to the right) in conditions '"One" access 1' and '"Two" access two'.

Testing the quantifier ‘some’

Next, we conducted a replication of the experiment with the use of the quantifier some (‘niektóre’). The participants had to bet again on zero, one, two and three objects having the property. In the original experiment, the condition ‘some access 2’ generated higher bets on three objects having the property rather than one object having the property. The replicated results provide a reverse trend, which points out to a lack of cancellation of the scalar implicature ‘not all’. In other words, the implicature was inferred, while the predictions were that it would not be formed (see Fig 5).

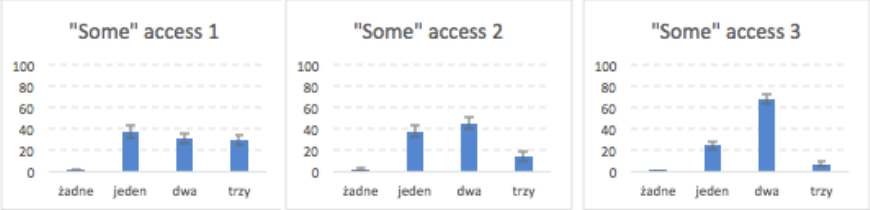

Fig 5 replication with the participation of 50 students

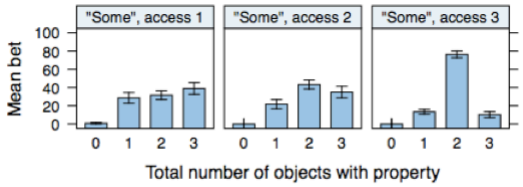

Fig 6 The original experiment in condition ‘”some” access two’. Reuse license number 4343821165053, source: https://onlinelibrary.wiley.com/doi/abs/10.1111/tops.12007

The means and standard deviations for the first experiment with the quantifier ‘some’ are presented in Figure 7.

	Original	Replication
Access 1	27.1	39,18
	(4,9)	(4,9)
Access 2	*34,8*	46,68
	(5,7)	(6,1)
Access 3	*93,0*	89,2
	(2,7)	(3,3)

Fig 7 Means and standard deviations for the replication with the quantifier ‘some’

Discussion

Our hypothesis was that a quantifier scope ambiguity occurred during the experiment, which is presented in Figure 8.

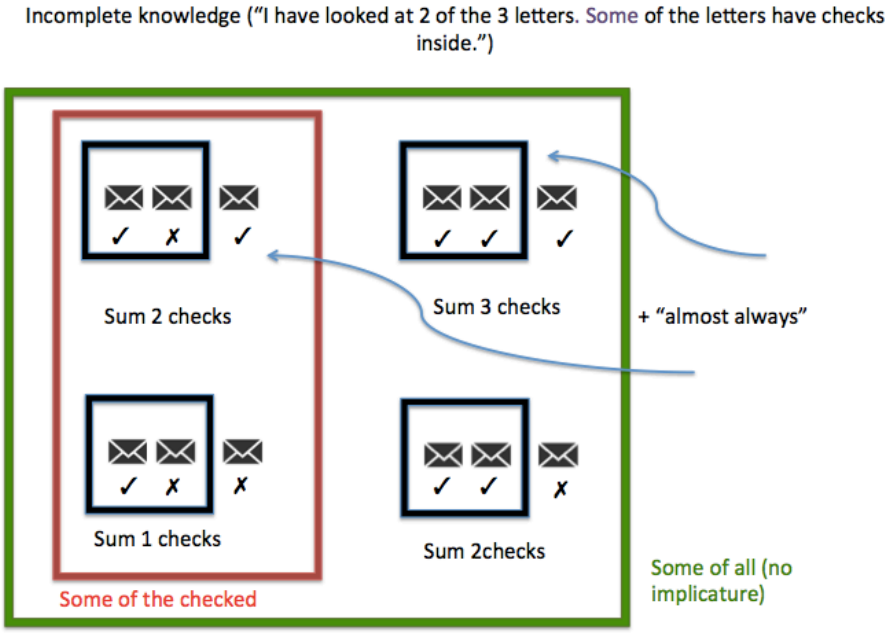

Fig 8 Hypothesis concerning quantifier scope ambiguity in the condition '"Some" access two'

Polish participants seemed not to interpret the speaker as indicating that some of all of the objects have the salient feature. Rather, they seemed to adopt the interpretation that some of the checked objects (for instance letters) had the salient feature (for instance checks inside). Let us focus on the '"Some" access 2' condition presented in Figure 5. The difference between the bets on one of the object having the property and two of the objects having the property are of 7,7%. This is an effect size of only 0,26%. By contrast, the difference between the bets on two of the objects having the property and three of the objects having the property are of 31,4%. This is an effect size of 1,12%. Thus, the bets on three are significantly lower than the bets on two, which supports the hypothesis of Polish speakers interpreting the scenario utterance as 'some of the checked objects'. We decided to test this hypothesis.

3 The quantifier scope disambiguation experiment

To test our hypothesis, we modified the content of the scenarios by adding 'some of all'. Thus, for instance the letter scenario took the following form:

'Letters to Laura almost always have checks inside.
Today Laura received three letters.
How many of the 3 letters do you think have checks inside?

Laura tells you on the phone: "I have looked at 2 of the 3 letters. ***Some of all*** of the letters have checks inside."
Now how many of the 3 letters do you think have checks inside?
Do you think Laura knows exactly how many of the 3 letters have checks inside?'

The results received (see fig 9) provided an only 7% weaker implicature that however still persisted since the bets on three objects having the property in questions were still lower than the bets on one object having the property in question compared to the previous results without introducing the disambiguation to the scenario (see Fig 5). In other words there was still no cancellation of the scalar implicature.

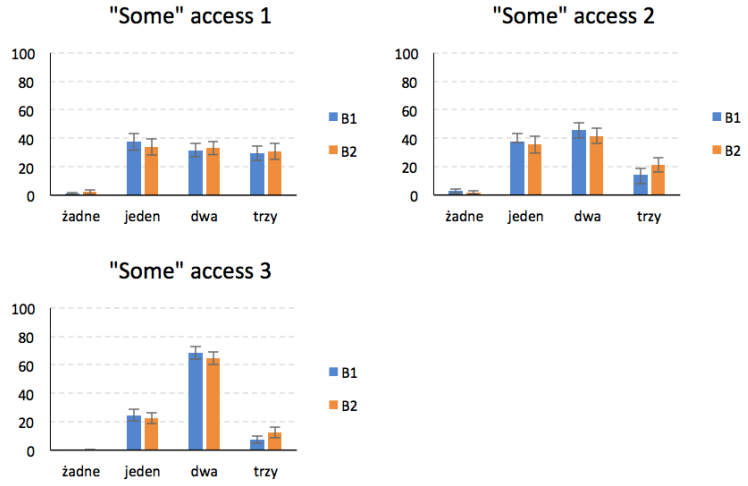

Fig 9 Comparing the 'some' quantifier experiment without disambiguation (B1) and with disambiguation (B2)

4 General discussion

Thus, the experiment depicted in Figure 9 disproved our hypothesis concerning quantifier scope ambiguity in experiments employing the quantifier 'some'. For this reason, there remained the question on why Polish participants did not cancel the scalar implicature 'not all' in partial knowledge conditions.

Our hypothesis is that Polish people are 'cache'ing' the mechanisms of implicature formation from full-knowledge contexts and are cognitively biased in applying them to contexts where the knowledge about the situation is incomplete.

What could back this hypothesis is the fact that the participants' bets while answering to the question measuring their prior probability distribution was not as predicted. Participants did not bet as if the relevant objects 'almost always' had the property in question. Thus, they probably ignored the 'almost always'- experimental manipulation in their answers on prior probability distribution and employed a prior probability distribution based on their every-day experience.

5 Conclusion

The carried experiments depicted that the RSA model is cross-linguistic as the patterns of interaction between knowledge and linguistic inference replicated. However, it also depicted differences in pragmatic inferences such as scalar implicatures carried by English and Polish native speakers.

6 References

1. Bergen L., Levy R., Goodman N., 2015, Pargmatic reasoning through semantic inference, Semantics and Pragmatics 9, doi:http://dx.doi.org/10.3765/sp.9.20
2. Benz, A., Jasinskaja K., 2017. Questions Under Discussion: From Sentence to Discourse, Discourse Processes, 54:3, 177-186, DOI: 10.1080/0163853X.2017.1316038
3. Chierchia, G., 2004. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface, in: Bettelli, A. (Ed.), Structures and beyond: The Cartography of Syntactic Structures. Oxford University Press, pp. 39–103.
4. Goodman, N.D., Stuhlmüller, A., 2013. Knowledge and Implicature: Modeling Language Understanding as Social Cognition. Top. Cogn. Sci. 5, 173–184. doi:10.1111/tops.12007
5. Grice, P., 1975. Logic and Conversation, in: Cole, P., Morgan, J.. (Eds.), Syntax and Semantics. Academic Press, New York.
6. Horn, L.R., 2006. Implicature, in: The Handbook of Pragmatics. Blackwell, Malden, MA; Oxford.
7. Noveck, I.A., Sperber, D., 2004. Experimental pragmatics. Palgrave Macmillan, Basingstoke, Hampshire; New York.

7 Appendix - vignettes

Link to repository: http://www.apohllo.pl

Below are the scenarios we employed translated into Polish:

1. Nasiona

Nasiona słonecznika prawie zawsze kiełkują w ciągu jednego dnia, gdy umieści się je w wodzie.
Dwa dni temu botanik Jan umieścił trzy nasiona słonecznika w wodzie.
Jak myślisz, ile ziaren słonecznika spośród trzech wykiełkowało?
Jan informuje Cię przez telefon: “Sprawdziłem dwa spośród trzech nasion.
Dwa nasiona wykiełkowały.”
Jak myślisz, ile spośród trzech nasion wykiełkowało?
Czy uważasz, że Jan wie ile dokładnie spośród trzech nasion wykiełkowało?

2. Losy

Losy na loterii fantowej prawie zawsze wygrywają.
Janusz zakupił trzy losy na loterii fantowej wczoraj.

Jak myślisz, ile spośród tych trzech losów wygrało?
Janusz informuje Cię przez telefon: “Sprawdziłem dwa spośród trzech losów.
Dwa spośród trzech losów wygrały.’
Jak myślisz ile spośród trzech losów wygrało?
Czy uważasz, że Janusz wie ile dokładnie spośród trzech losów wygrało?

3. Egzaminy

Uczestnicy kursu ‘Wprowadzenie do biologii’ prawie zawsze zdają egzamin.
Trzech studentów Marka podeszło wczoraj do egzaminu z ‘Wprowadzenia do biologii’.
Jak myślisz, ilu spośród tych trzech studentów zdało egzamin?
Marek informuje Cię przez telefon: “Sprawdziłem dwa spośród trzech egzaminów.
Dwóch studentów zdało egzamin.”
Jak myślisz, ilu spośród trzech studentów zdało?
Czy uważasz, że Marek wie ilu dokładnie spośród trzech studentów zdało egzamin?

4. Owoce

Jagody Goi to małe owoce, które prawie zawsze mają we wnętrzu wyschnięty miąższ.
Monika zakupiła trzy jagody Goi wczoraj.
Jak myślisz, ile spośród trzech jagód Goi ma wyschnięty miąższ?
Monika informuje Cię przez telefon: “Sprawdziłam dwa spośród trzech owoców.”
Dwa owoce mają wyschnięty miąższ.
Jak myślisz, ile spośród trzech owoców ma wyschnięty miąższ?
Czy uważasz, że Monika wie ile dokładnie spośród trzech owoców ma wyschnięty miąższ?

5. Telefony

Zepsute telefony marki Sigo prawie zawsze mają przepalone tranzystory.
Bartosz musi naprawić trzy telefony marki Sigo.
Jak myślisz, ile spośród trzech telefonów marki Sigo ma przepalone tranzystory?
Bartosz informuje Cię przez telefon: “Sprawdziłem dwa spośród trzech telefonów.”
Dwa telefony mają przepalone tranzystory.
Jak myślisz, ile spośród trzech telefonów ma przepalone tranzystory?
Czy uważasz, że Bartosz wie ile dokładnie spośród trzech telefonów ma przepalone tranzystory?

6. Listy

Listy do Laury prawie zawsze mają czeki w środku. Dzisiaj Laura otrzymała trzy listy.
Jak myślisz, ile spośród trzech listów ma czeki w środku?
Laura informuje Cię przez telefon: “Sprawdziłam dwa spośród trzech listów.
Dwa listy mają czeki w środku.”
Jak myślisz, ile spośród trzech listów ma czeki w środku?
Czy uważasz, że Laura wie ile dokładnie spośród trzech listów ma czeki w środku?