On the interpretation of German *einige*. The effect of tense and cardinality

Maya Cortez Espinoza & Lea Fricke*

Abstract. We present a study investigating the effect of tense (past vs. future) on the computation of scalar implicatures in connection with the German quantifier *einige* 'some' in an interactive experiment, which included a financial incentive for participants to consider whether another speaker would share their judgment. We tested the hypothesis that scalar implicatures are less frequently drawn in future tense than in past tense. In addition, we studied to what extent sets with various cardinalities are prototypical representatives of *einige* + *N*. We hypothesized that larger cardinalities are more prototypical representatives of the quantifier *einige* than smaller cardinalities (relative to the cardinality of the total set). We analyzed the experimental data with probabilistic Bayesian models with a linking hypothesis between participants' responses and readings based on utility maximization in simple decision problems. In line with the hypotheses, we found that less scalar implicatures are drawn in future tense than in past tense, which replicates the results of previous research on English *some*, and that with an increase in set size acceptance of statements involving *einige* also increases.

Keywords. Scalar implicatures; German *einige*; tense; quantifier interpretation; Bayesian Modelling; experimental pragmatics;

1. Introduction. A sentence like (1-a), involving the quantifier *some*, usually triggers the scalar implicature (SI) (1-b).

(1) a. Some students danced.
b. SI: Not all students danced.

A factor that influences the computation of SIs, but that has not yet received much attention in the literature, is tense, to the effect that SIs are less frequently drawn in future tense than in past tense. Normally, the Gricean reasoning process upon hearing a statement involving the quantifier *some*, like (1-a), goes roughly as follows: Under the assumption that the speaker is cooperative and follows the maxim of quantity, a hearer assumes that she chose *some* as the maximally informative expression and that she cannot truthfully use the stronger expression *all*. Thus, the hearer concludes that not all students danced.[1]

Moreover, as observed by Geurts (2010), the assumption that the speaker is maximally informed about the issue at hand ("competence assumption") is crucial for drawing the implicature. Not being able to truthfully assert the stronger alternative *All students danced* is compatible with

*We would like to thank Edgar Onea, Daniele Panizza, Swantje Tönnis, and Alexandre Cremers for their insights in fruitful disscusions and helpful suggestions. We further thank Simon Dampfhofer, Kathrin Hirsch, Kassandra Keinz, Lea-Sophie Kravanja, Michaela Leeb, Melanie Loitzl, and Eva Winkler for their help conducting this experiment. This research was funded by LingLab Graz, which we thankfully acknowlegde. Authors: Maya Cortez Espinoza, Karl-Franzens-Universität Graz (maya.cortez-espinoza@uni-graz.at) & Lea Fricke, Karl-Franzens-Universität Graz/Ruhr-Universität Bochum (lea.fricke@rub.de).

[1] Alternatively, SIs can be derived from applying a silent logical operator on various sites (Fox 2007, Chierchia 2013). With this account, SIs can be derived both at the utterance level and in embedded positions.

not being opinionated about the stronger alternative. Thus, to take the "epistemic step" (Sauerland 2004) and conclude that the speaker implicated that not all students danced, the hearer needs to assume that the speaker is competent about the issue.

As the future is inherently uncertain, a statement about the future, like (2) is a mere prediction. For this reason, it makes sense for the speaker to use a weak expression to make her statement semantically compatible with both outcomes (*Some or all students danced*). On considering this, the hearer will not take the epistemic step and no implicature will be drawn.

(2) Some students will dance.

This effect of tense has to our knowledge only once been investigated for English and Italian in an acquisition experiment by Chierchia et al. (1998)[2]. The authors argue that a prediction about the future constitutes a context in which an SI is suspended. In the experiment, children had to judge the truth of statements made by a puppet, which contained one of the scalar expressions *some* and *or*. Two contexts were tested. In the first context, which Chierchia et al. (1998) label 'description mode', the target sentence is in past tense and in the second, which is called 'prediction mode', it is in future tense. With regard to *some*, they report preliminary results according to which there was 14% and 34% acceptance for the SI violation in English and Italian respectively in the description mode and more than 75% acceptance in both languages in the prediction mode. The final results for *or* are similar: In the description mode, the acceptance rate for the SI-violation is lower than 33% in both languages while it is at ceiling in the prediction mode. The results of this study support the hypothesis that tense affects the computation of SIs and that SIs are less frequently drawn in future tense than in past tense. However, its limitations are that there are no data for adult speakers and that in the case of *some* only preliminary data are reported. We aim to replicate these results for German *einige* 'some' with adult speakers.

For the interpretation of the quantifier *einige* or *some*, another factor beside the SIs comes into play, namely the question which cardinalities can be felicitously described with this quantifier. According to the set-theoretic definition, a statement involving *some*, like (1-a), is true iff the intersection between the set of dancers and the set of students is non-empty (e.g. Barwise & Cooper 1981). However, previous research suggests that the exact cardinality of the intersection matters to people, i.e. some cardinalities are more typical representatives of *some* than others.

Van Tiel & Geurts (2014) studied the interpretation of quantifiers in English. They presented participants with pictures of ten circles, of which varying numbers were black, together with the sentence ***Quantifier*** *of the circles are black*. Participants had to either judge the truth value of the statement or rate on a Likert scale how well the statement describes the situation (typicality). They found that for the quantifier 'some', the numbers 2–4 received the highest truth ratings while the numbers 4–6 received the highest typicality rating.

Another study investigating the processing of *some*, Degen & Tanenhaus (2015), included naturalness ratings of a statement like "You got some of the gumballs" with contexts of varying numbers from a total of 13 gumballs. Additionally to judging naturalness, participants could judge a statement as false. The authors found that statements in which *some* referred to the set sizes 1, 2, 3 or 13 gumballs were rated as less natural than statements in which *some* referred to sets of 6–8

[2]We thank an anonymous reviewer for pointing out this paper to us.

gumballs. When *some* referred to a set of one gumball, the statement was rated as false in 12% of the cases. They also found that particularly for small but also for large set sizes, exact number terms were rated as more natural than *some*.

We note that in both studies numbers that roughly represent half of the total sample receive the highest naturalness/typicality ratings and that it appears to be odd to refer to a set size of 1 with *some*. Responsible for the latter effect seems to be the plural marking of the noun phrase which leads to a plurality inference, as discussed in Spector (2007).

In our study, we compare the acceptability of statements involving *einige* referring to different set sizes ranging from 1–9. Note that German has two different versions of 'some', *manche* and *einige*. According to our intuitions, *manche* is associated with smaller numbers compared to *einige*. Additionally, when stressed *einige* has a prominent reading of 'many', as example (3) shows.

(3) A: Wie viele Leute haben bereits gegessen? ('How many people have eaten already'?)
B: Naja, schon EINIGE. ('Well, QUITE a few'.)

Thus, we want to test the hypothesis that statements involving *einige* will receive higher acceptance when referring to sets with higher cardinalities, relative to the cardinality of the total set, than when referring to sets with smaller cardinalites.

A further goal of our experiment was to control for the kind of reasoning induced by the experimental task. Experiments with a classic truth-value judgement task design often only test for the general availability of a reading but do not control for the status of that interpretation with the consequence that it is not clear whether participants' judgments are based on a semantic interpretation or whether other factors, such as communicative relevance, were considered. Previous research has shown that the nature of the task crucially affects the rate of drawing implicatures. For example, with regard to embedded implicatures, Geurts & Pouscoulous (2009) showed that an inference task lead to a higher implicature rate compared to a truth-value-judgement task. Benz & Gotzner (2014) suggested that relevance of the SI in the context of the experiment is a decisive factor for its computation.

In our experiment, we employed the method developed by Fricke et al. (2022). It aims to target interpretations that are of communicative relevance, that is, interpretations that the participant deems likely to be shared by another language user. To induce this kind of recursive thinking in participants, the design includes a monetary incentive: Participants were told that they would lose money if another person did not share their judgment. This way, every single response had direct financial consequences for the participant, and it was in their own interest to think carefully about each item. The linking hypothesis is based on utility theory for simple decision problems: Participants aim to maximize their expected utility measured in terms of financial payoff.[3]

2. Experiment.

We investigate the following hypotheses in our experiment:

- The computation of SIs of German *einige* is affected differently by future tense and past tense to the effect that the rate of implicatures is higher in past tense.

[3]We are aware that utility theory has received criticism, e.g. by Tversky (1975) and Kahneman & Tversky (1979). However, in our experiment the translation between decisions and financial benefits is quite simple. Moreover, we introduced some level of control regarding the effect of bias in participants' decisions and found no such effect.

- Larger cardinalities are more prototypical representatives of the quantifier *einige* than smaller numbers and thus statements involving the quantifier *einige* will receive higher acceptance rates when referring to larger sets than to smaller sets.
- The set of cardinality 1 is a particularly bad representative of the quantifier *einige* due to the plurality inference.

2.1. PARTICIPANTS. We tested 32 native speakers of German (mostly Austrian German) (mean age = 23.8 years, SD = 5.5 years, 15 female and 17 male), all of which were university students or former university students. They were recruited via a university-newsletter and received a financial compensation varying between 8.50€ and 11.20€, with a mean of 10.18 and a standard deviation of 0.56, depending on their performance on control items.

2.2. MATERIALS. The target sentences were conditional statements containing the scalar term *einige*. They were presented in the context of a story about nine candidates in a reality show, who did activities together. The stimuli had the form of bets, made by a person named Lina about activities to happen on the show and the participants' task was to decide whether the bets were won (called 'accepting the bet' in the following) or lost ('rejecting the bet'), thereby judging the truth of the target sentences. We manipulated three factors.

The first factor was CARDINALITY, which represents the number of candidates that were involved in an activity and which ranged from 0 to 9. The 0-context yields a false target sentence, the 9-context yields an SI-violation and the numbers 1–8 constitute different manifestations of the quantifier *einige*. CARDINALITY was tested within participants and within items.

The second factor was TENSE. The verb of the target sentences was either in past tense (*Perfekt*)[4] or in future tense (*Futur I*). This factor was tested at the participant level. Participants were assigned either past tense or future tense. (Half of the participants saw bets in past tense and the other half in future tense.) For past tense, people were told that the show had been prerecorded already before Lina placed the bets but aired later, and therefore, Lina worded her bets in past tense. For future tense, Lina placed her bets before the recording of the show and therefore, bets were worded in future tense.

The third factor, ROLE, was manipulated at the participant level as well (meaning half of the participants acted in role 1 and the other half in role 2). In **role 1**, participants had to decide whether or not to redeem bets at a betting office. They received 5€ starter cash. Redeeming a bet cost a fee of 10 cents each. For each accepted bet that was actually won the participants received 30 cents payout. Thus, in effect, the participants gained 20 cents for an accepted bet which was won, and they lost 10 cents for an accepted bet that was lost. In this role, a participant profited financially from accepted bets. To avoid a bias towards accepting borderline cases in this role, we installed the redeeming fee, with which participants would lose money when randomly submitting bets and would therefore consider whether a betting office agent might share their judgment. In **role 2**, participants acted as a betting office agent. They had to decide for each redeemed bet whether it was won or not. Participants received 15€ starter cash. For an accepted bet, they had to pay out 20 cents, while their budget remained unchanged when they rejected a bet. Also, participants were told that Lina, the person who had placed the bets, would raise an objection if a bet that was in fact

[4]In colloquial German, perfect tense is the default tense to talk about the past, see e.g. Klis et al. (2017).

won had been rejected. This would lead to a monetary loss of 30 cents. Thus, participants had to consider whether their judgement would be shared by Lina. In sum, in role 2, participants profited from rejected bets. To avoid a bias towards rejecting bets, we installed the financial deduction for incorrect decisions. These conditions for the two roles are summarized in Table 1.

Role	Status of bet	Accepted bet	Rejected bet
Role 1	In fact won	+20ct	0ct
	In fact lost	-10ct	0ct
Role 2	In fact won	-20ct	-30ct
	In fact lost	-20ct	0ct

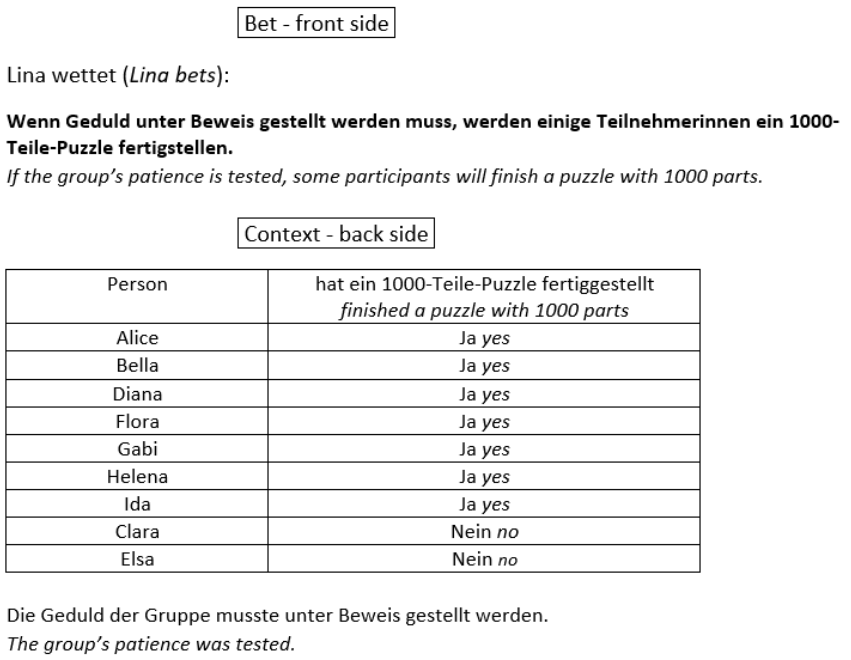
Table 1: Payoff table

Note that in calculating the compensation that a participant received, decisions on test items were always considered correct; only mistakes on fillers affected the final compensation negatively. However, participants did not know this before the experiment.

To sum up, the factorial design was 10 CARDINALITY x 2 TENSE x 2 ROLE. There were two lexicalizations per cardinality, which varied between lists. Test items were distributed over 16 experimental lists (half of them in future, half of them in past tense) with 20 test items and 2 participants per list. In addition to the test items, each list contained 32 filler items. 12 of of them were test items from a different experiment, and 9 uncontroversially won and 9 uncontroversially lost bets served as controls.

Figure 1 is an example of a test item for future tense. The stimuli were shown along with a table which indicated for each candidate whether she was involved in the activity in question. The numbers of involved candidates ranged from 0 to 9 (CARDINALITY). Additional context resolved the antecedent of the conditional as true. Note that target sentences had the form of conditional sentences because we had an additional hypothesis about upward and downward entailing environments that we will not report on due to a procedure error, which happened on the level of participant instructions and which turned part of the data invalid.

2.3. PROCEDURE. The experiment took place in the lab of the theoretical and empirical linguistics research group at the University of Graz. Before the start of the actual experiment, participants saw 5 training items. They a) helped to get the participants accustomed to the task and b) made clear how to handle the conditional construction, which included the target sentence. Then, depending on the role they were assigned, participants each received 5 or 15€ starter cash in stacks of 10 and 20 cent coins. The participant saw the betting slips one at a time in a randomized order and had to decide whether to accept (pay out/redeem) the bet or not. If they wanted to accept, they had to return the betting slip together with the money to the experimenter. The experimenter entered the participant's decision into an excel sheet that automatically calculated the sum the participant received as financial compensation after the experiment. The participant did not receive any feedback as to their gains and losses from individual bets, neither during nor after the experiment. After the first half of the betting slips, there was a short break. The experiment took 25 to 40 minutes in total.



Bet - front side

Lina wettet (*Lina bets*):

Wenn Geduld unter Beweis gestellt werden muss, werden einige Teilnehmerinnen ein 1000-Teile-Puzzle fertigstellen.

If the group's patience is tested, some participants will finish a puzzle with 1000 parts.

Context - back side

Person	hat ein 1000-Teile-Puzzle fertiggestellt *finished a puzzle with 1000 parts*
Alice	Ja *yes*
Bella	Ja *yes*
Diana	Ja *yes*
Flora	Ja *yes*
Gabi	Ja *yes*
Helena	Ja *yes*
Ida	Ja *yes*
Clara	Nein *no*
Elsa	Nein *no*

Die Geduld der Gruppe musste unter Beweis gestellt werden.

The group's patience was tested.

Figure 1: Sample item, cardinality 7, future tense

2.4. Results.

2.4.1. Data exclusion. The data of one participant (role 1, past tense) was excluded due to erring three out of 18 times on control items. Therefore, data from 31 participants was subjected to further analysis (7 for the combination of role 1 and past tense, and 8 for all other TENSE x ROLE combinations).

2.4.2. Descriptive statistics. Figure 2 shows the acceptance rates of bets by CARDINALITY and TENSE. In the cardinality-0 context, in which no candidate was involved in the activity, acceptance is at 0 in both tenses. Acceptance in the cardinality-9 context, which constitutes a SI violation, differs between the tenses. In the past tense, acceptance is at 60% while it is at 93.8% in future tense. Turning to the acceptance rates for cardinalities that represent *einige* 'some', we observe a) acceptance for cardinality 1 is particularly low (0/13.3%) and b) that acceptance increases with increasing cardinality. Moreover, acceptance rates for cardinalities 1–4 differ between the tenses; they are higher in past tense than in future tense. Our descriptive data shows that the two roles hardly differ in acceptance rates, as can be seen in Table 2. Therefore, we did not consider the factor ROLE in the Bayesian modelling.

Role	Won: count	Lost: count
1	108 (72.0%)	42 (28.0%)
2	115 (71.9%)	45 (28.1%)

Table 2: Accepting rates between roles

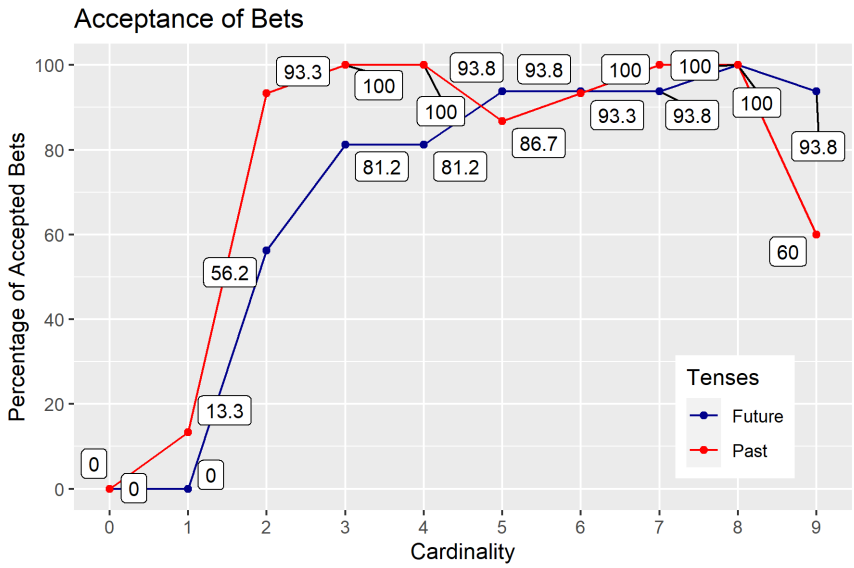


Figure 2: Percentage of accepted bet by tense and cardinality

3. Bayesian Modelling.

3.1. CONCEPT. We created a Bayesian model based on the assumption that a participant aims to maximize her utility for each decision in the experiment. In this model, a participant's expected utility to act (u_{act}), shown in (4) is the sum of the products of i) the subjective probability they assign the decision to be right ($PR^s_{act,\,true}$) and the payoff for being right and ii) the subjective probability they assign to the decision being wrong ($PR^s_{act,\,false}$) and the payoff for being wrong. $PR^s_{act,\,true}$ is different for the act of *accepting* and *rejecting*, since $PR^s_{reject,\,true} = 1 - PR^s_{accept,\,true}$.

(4) $$u_{act} = \text{payoff}_{act,\,true} * PR^s_{act,true} + \text{payoff}_{act,\,false} * (1 - PR^s_{act,\,true})$$

The probability for accepting a bet is then determined from the overall utility (the difference between u_{accept} and u_{reject}) as log odds:

(5) $$p(accept) = logistic((u_{accept} - u_{reject}))$$

Positive overall utilities, therefore, mean higher than 50% chance of accepting bets, negative utilities mean lower than 50% of accepting bets and utilities further away from 0 mean stronger tendencies to accept/reject. Intuitively, therefore, strong PR^s and big rewards/punishments influence u greater than weak PR^s and small rewards/punishments. In our model, we do not assume differences in PR^s for different people as we did not have any hypotheses on some people acting differently than others.

With the utility formula being constant across conditions, only subjective probabilities alternate. These conceptually depend on two factors: the computation of an SI and the sensitivity of *einige* to different cardinalities. Firstly, we assumed a general probability $PR(SI)$ for interpreting *einige* with the SI *not all*. The alternative, $PR(noSI)$, is the probability for interpreting *einige* without the SI. As there are no other competing readings of *einige*, we assumed $PR(noSI) = 1 - PR(SI)$. Our model assumes this probability to be equally accessible to all

speakers of German. Furthermore, our model assumes these probabilities to be different for the two tense levels: The probability to interpret *einige* without the SI is higher in future tense than in past tense.

Of the two probabilities $PR(SI)$ and $PR(noSI)$ either both, only one, or none of them contributes to $PR^s_{accept,\ true}$, depending on whether the context supports the respective reading, see Table 3. A participant's $PR^s_{accept,\ true}$ is therefore the sum of her commitments to supported readings. This means that for cardinality 0, where none of the readings are supported, $PR^s_{accept,\ true}$ is set to 0, irrespective of the probabilities of the two readings. For cardinality 9, $PR^s_{accept,\ true}$ depends solely on $PR(noSI)$, which enables us to tease apart the two readings' probabilities.

Cardinality	NoSI-reading	SI-reading
0	not supported	not supported
1–8	supported	supported
9	supported	not supported

Table 3: Possible contexts and supported readings

The second factor that contributed to $PR^s_{act,\ true}$ is cardinality sensitivity. For cardinalities 1–8, both the noSI- and the SI-reading are supported and $PR^s_{accept,\ true}$ yields 1, but in these cases, cardinality sensitivity accounts for lower accepting rates.

In our model, cardinality sensitivity is represented by a Dirichlet distribution over cardinalities 1–8. The values of prototypicality are determined by the model itself through sampling. Conceptually, we were interested in the prototypicality of all cardinalities relative to each other with the most prototypical cardinality having $PR^s_{accept,\ true}$ = 1. For this, each prototypicality value is divided by the highest prototypicality value. To sum up, Table 4 shows all the possible cases for calculating $PR^s_{accept,\ true}$. They depend on cardinality (c) and tense (t).

Case	Calculation of $PR^s_{accept,\ true}$
if c = 0:	$= 0$
else if $1 \leq$ c ≤ 8:	$= \frac{prototypicality(c)}{max(prototypicality)}$
else if c = 9 and t = past:	$= PR(noSI)_{past}$
else (c = 9 and t = future):	$= PR(noSI)_{future}$

Table 4: Calculating $PR^s_{accept,\ true}$

3.2. Implementation and Output. We conducted a MCMC Simulation to construct our Bayesian Model in STAN using the RStan package in R (Stan Development Team 2022b,a, R Core Team 2022). We used uniform priors in order not to skew our results with previous hypotheses. When running the model, we used 15000 iterations, the first 5000 of which were dismissed as warm up. All models had excellent convergence as evidenced by visual checks and Rhat values (1.00–1.01). After sampling, means were computed for each desired parameter. For model comparison, we used the bridgesampling package (Gronau et al. 2020) and computed Bayes Factors

(BF) for samples of 8000 iterations each.

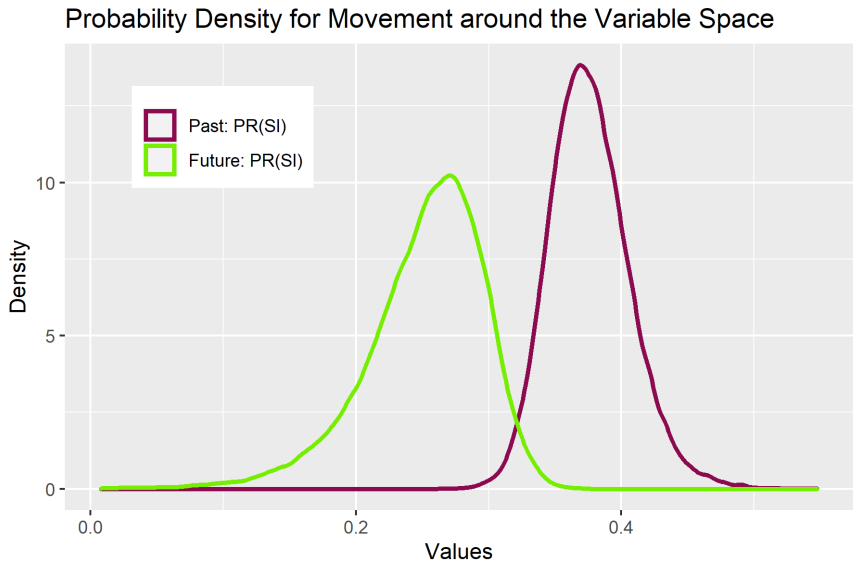


Figure 3: Probability density plot for $PR(SI)$ (with $PR(noSI)$ = 1-$PR(SI)$)

As shown in Table 5, the probability to interpret *einige* without the SI is 13% higher for future tense than for past tense.

	$PR(SI)$	$PR(noSI)$
past	0.38	0.62
future	0.25	0.75

Table 5: SI probabilities

We additionally created a reduced model that does not distinguish between reading probabilities for tenses and compared it to our original model, which turned out superior to the reduced model (BF = 5.15).

Figure 3 shows a probability density plot for the movements of the parameter around the variable space during the simulation. The x-axis shows different probability values. For prototypicality values, see Table 6 and the associated probability density plot in Figure 4. Cardinality 1 has the lowest prototypicality value with 0.05. Cardinalities 2–7 are quite similar to each other with values ranging from 0.12 – 0.14, and cardinality 8 has the highest prototypicality value with 0.18. Model comparison with a reduced model, in which no cardinality relative prototypicality was assumed, yielded that our original model is superior (BF $>$ 150).

Cardinality	1	2	3	4	5	6	7	8
Prototypicality	0.05	0.12	0.13	0.13	0.13	0.13	0.14	0.18

Table 6: Prototypicality values

As our descriptive data suggested that there might be a difference for cardinality sensitivity between tenses (small numbers received less acceptance in future tense than in past tense), we

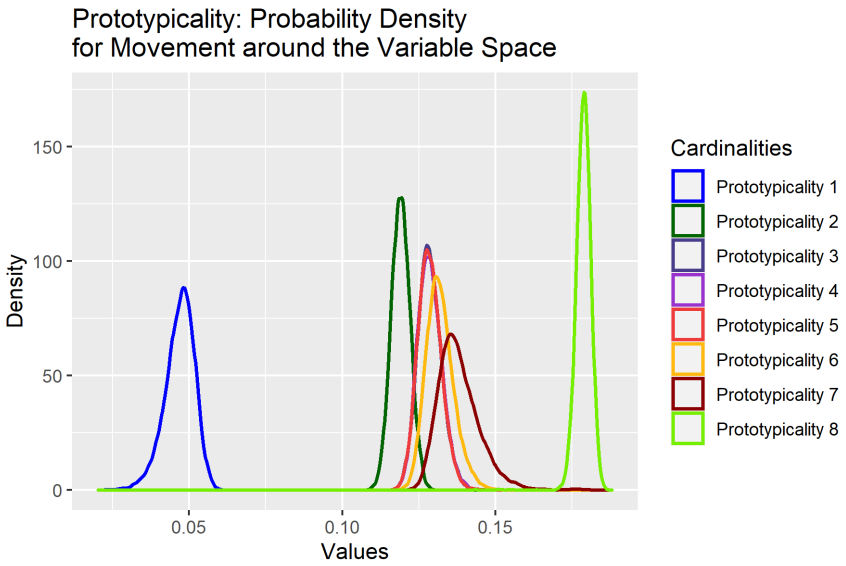


Figure 4: Probability density plot for prototypicality values

constructed another model that specified two different Dirichlet distributions for prototypicality (future vs. past tense).[5] Model comparison showed that the original model was better than this more complex model (Bayes Factor $>$ 150).

4. Discussion. We found that SIs are drawn less reliably in future tense than in past tense. The method we employed was intended to evoke rational behavior and targeted interpretations that are relevant from a communicative perspective. Replicating the findings by Chierchia et al. (1998) with this method suggests that the SI-difference caused by tense that was found in that study is a robust effect and holds in communicative settings.

As discussed in the introduction, this phenomenon can be explained pragmatically, namely with the epistemic step not being possible for future situations. However, in the case of our experiment, this explanation is not completely conclusive. In our design, the context was exactly the same for future tense and past tense. In both cases, the bet was made about an unknown situation. Still, a clear difference was found between future tense and past tense. This hints at the effect of tense on SIs being at least partly in the semantic domain – possibly the effect is a grammaticalized instance of the pragmatic principle described above.

A further finding of our experiment is that the acceptability of statements involving *einige* increases with growing set sizes. This deviates from the findings in the literature on English *some*, according to which medium set sizes receive the highest ratings. This difference may stem from the fact that *einige* competes with another lexical item, *manche*, which is according to our intuitions associated with small numbers, at least in some cases.[6] Furthermore, stressed *einige* has a very

[5]Note that we did not have a hypothesis about this. So this is a post hoc analysis.

[6]For *manche* to be felicitous, a basic set must be specified, which is not the case for *einige*. Compare (i-a) in an out of the blue-context:

(i) a. Heute um sechs Uhr haben sich einige/#manche Leute am Hasnerplatz verabredet.

prominent reading of *many*, which is clearly not the case for English *some*.

Although, descriptively, we observed a tendency in the data for lower cardinalities to be less accepted in future tense than in past tense, this difference was not meaningful in the Bayesian analysis. Nevertheless, it would be interesting to gather more data, as an experiment targeted at this issue may find this to be significant.

Another idea for future research is testing whether there is a contrast between the rates at which SIs are drawn between the English *going-to*-future and *will*-future. As the former is used to refer to future events that are more certain than the latter, we expect the future effect to be smaller for the *going-to*-future and the rate of drawn SIs to be higher.

References

Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2). 159–219. https://doi.org/10.1007/BF00350139.

Benz, Anton & Nicole Gotzner. 2014. Embedded implicatures revisited: Issues with the truth value judgement paradigm. In Judith Degen, Michael Franke & Noah Goodman (eds.), *Proceedings of the formal & experimental pragmatics workshop*, 1–6. Tübingen.

Chierchia, Gennaro. 2013. *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199697977.001.0001.

Chierchia, Gennaro, Stephen Crain, Maria Teresa Guasti & Rosalind Thornton. 1998. "some" and "or": A study on the emergence of logical form. In *Proceedings of the 22nd annual boston university conference on language development (bucld)*, vol. 22, 97–108.

Degen, Judith & Michael K. Tanenhaus. 2015. Processing scalar implicature: A constraint-based approach. *Cognitive Science* 39. 667–710. https://doi.org/10.1111/cogs.12171.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. *Presupposition and Implicature in Compositional Semantics* 71–120. https://doi.org/10.1057/9780230210752$_4$.

Fricke, Lea, Emilie Destruel, Malte Zimmermann & Edgar Onea. 2022. The pragmatics of exhaustivity in embedded questions: An experimental comparison of *know* and *predict* in German. submitted.

Geurts, Bart. 2010. *Quantity implicatures*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511975158.

Geurts, Bart & Nausicaa Pouscoulous. 2009. Embedded implicatures?!? *Semantics & Pragmatics* 2. 1–34. https://doi.org/10.3765/sp.2.4.

Gronau, Quentin F., Henrik Singmann & Eric-Jan Wagenmakers. 2020. bridgesampling: An R package for estimating normalizing constants. *Journal of Statistical Software* 92(10). 1–29. https://doi.org/10.18637/jss.v092.i10.

Kahneman, Daniel & Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47(2). 263–291. https://doi.org/10.2307/1914185.

Klis, Martijn van der, Bert Le Bruyn & Henriette De Swart. 2017. Mapping the perfect via translation mining. In *Proceedings of the 15th conference of the european chapter of the association for computational linguistics*, vol. 2, 497–502. https://aclanthology.org/E17-2080.

'Today at six, einige/#manche people arranged to meet at Hasnerplatz.'

R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27. 367–391. https://doi.org/10.1023/B:LING.0000023378.71748.db.

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: on higher order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, 243–281. Houndsmills et al.: Palgrave-Macmillan. https://doi.org/10.1057/9780230210752_9.

Stan Development Team. 2022a. RStan: the R interface to Stan. R package version 2.21.5. https://mc-stan.org/.

Stan Development Team. 2022b. Stan modeling language users guide and reference manual, version 2.30.1. http://mc-stan.org/.

Tiel, Bob van & Bart Geurts. 2014. Truth and typicality in the interpretation of quantifiers. In *Proceedings of Sinn und Bedeutung (SuB)*, vol. 18, 451–468. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/327.

Tversky, Amos. 1975. A critique of expected utility theory: Descriptive and normative considerations. *Erkenntnis* 9(2). 163–173. https://www.jstor.org/stable/20010465.