ABSTRACTS
(in alphabetical order)
**Spatial distribution and exhaustivity requirement in Serbian**

1,2 Ana Bosnić and 2 Nataša Knežević

1 University of Groningen, 2 University of Nantes

Contact: a.bosnic@rug.nl

**BACKGROUND** Distributive universal quantifiers (e.g. English *each*) take as their restrictor the NP denoting what is distributed over (the distributive key in Choe’s 1987 terminology) and they involve distribution over *individuals*. In contrast, distributive markers, such as Serbian *po*, attach to the NP denoting what is being distributed (the distributive *share*) and it has been claimed they involve distribution over *spatiotemporal units* (cf. Balusu, 2006; Cable, 2014; Knežević, 2015).

The question at issue in the literature is whether the latter form of distribution involves universal quantification over *space/time units*. A standard diagnostic for universal quantification is exhaustivity (Mathewson, 2000; Zimmerman, 2002). Take the English sentence in (1) where the distributor *each* takes as its key (restrictor) *monkeys*. (1) is judged FALSE in the context depicted in the picture A1: the set of monkeys is not exhausted since there are non-jumping monkeys. Crucially, however, the Serbian sentence in (2) with the distributor-share marker *po* is judged TRUE with A1 – thus violating the exhaustivity requirement on participants.

(1) Each monkey jumped.

(2) Skače po jedan majmun.

‘One monkey jumps at different locations/time intervals.’

There are two accounts of this phenomenon. Balusu (2006) claims distribution over spatiotemporal units arises via universal quantification: the distributor takes as its key (restrictor) spatiotemporal units distributing events exhaustively over them. For Balusu (2) means that *in every relevant spatiotemporal unit, there is a monkey jumping.* (2) is thus correctly predicted to be true under A1 since all that is required is that quantification be exhaustive relative to spatiotemporal locations (i.e. cages), but not participants (i.e. monkeys). In contrast, on Knežević’s account, distribution over spatiotemporal locations does not arise via universal quantification but rather via event pluralization: (2) means that *there are at least two events of a monkey jumping in a cage.* (2) is also correctly predicted to be true for A1 since distribution need not be exhaustive, be it relative to participants or spatiotemporal locations.

In two Truth Value Judgment Task with Serbian adults we investigated whether distribution over spatiotemporal units is exhaustive and how spatiotemporal units are determined with *po*. We used variations of sentences like (2) and the five pictures (A1, A2, B, C1 and C2).

Balusu’s universal quantification analysis predicts that (2) should be true with pictures A1, A2 and C2. The predictions for B and C1 are false, because relevant spatial units do not contain jumping monkeys. Knežević’s account, on the other hand, predicts all conditions to be accepted, because they all satisfy at least two events of monkey-jumping in relevant spatial units.

**RESULTS EXPERIMENT 1** The acceptance of B suggested that spatial units do not need to be exhausted, supporting Knežević. However, the clear rejection of C1 show there is some kind of exhaustivity imposed. The question is whether this is exhaustivity over the pairings of
participants (monkeys) and spatial units (cages) or over the groups of participants. The follow-up was designed to answer this question using pictures A2, B and C2, where A2 and C2 now simply show groupings of animals, not bounded by explicit spatial units in the form of cages/caves.

RESULTS EXPERIMENT 2 C2 was rejected while A2 and B were considered true. This reveals that exhaustivity is required over the relevant sets of participants. Furthermore, the fact that participants did not simply restrict spatiotemporal locations to cages is surprising. Therefore, we hypothesize that a relevant spatial (or temporal) unit is determined by the relevant sets of participants, when the contextual method of division is not salient enough or it is not explicitly given. In other words, relevant spatiotemporal units are indeed the sorting key and they are determined by visual groupings of participants. Finally, these relevant sets of participants (monkeys) need to be exhausted by the jumping event (i.e. they must all include one jumping monkey). We will discuss these results, including recent results with the Korean distributive marker -ssik which show the same pattern as Serbian, as well as the theoretical implications for theories of distributivity in our talk.

REFERENCES


The timing of quantifier scope interpretation – Eyetracking evidence against incremental processing

Oliver Bott

Project Composition in Context, XPrag.de, University of Tübingen
(joint work with Fabian Schlotterbeck, Tübingen)

Contact: oliver.bott@uni-tuebingen.de

The talk presents two eyetracking studies investigating the online interpretation of quantifier scope. The first study investigated scope reconstruction in German doubly quantified sentences (Bott & Schlotterbeck, 2015). We tested sentences in which the first quantifier contained a variable jeden seiner Studenten (each of his students) that had to be bound by a second quantifier (exactly one professor) such as (1-a). Scope reconstruction was compared to a doubly quantified control condition (1-b) allowing for a linear interpretation. Across a number of eyetracking measures clear scope inversion costs were observed, but these costs only emerged late after participants had read the entire sentence (here, at gelobt). Thus, scope reconstruction seems to be a last resort only considered if no other option is available.

(1-a) Jeden seiner Studenten hat genau ein Professor… gelobt.
  each of his students,object has exactly one professor,subject… praised.
  ‘Exactly one professor praised each of his students…’

(1-b) Jeden dieser Studenten hat genau ein Professor… gelobt.
  each of these students,object has exactly one professor,subject… praised.
  ‘Exactly one professor praised each of these students…’

In a second study, we investigated the time course of linear scope construal in sentences with a quantifier and negation. Semantic complexity was manipulated by comparing monotone increasing (UE) and monotone decreasing (DE) quantifiers (cf. Deschamps et al., 2015) in interaction with the presence or absence of sentence negation. An offline pretest confirmed that the sentences were interpreted as intended and a first eyetracking experiment established clear processing differences between negated and non-negated scope disambiguated sentences with DE vs. UE quantifiers. DE quantifiers incurred overall more processing costs than UE quantifiers, and these processing costs interacted with the presence or absence of negation in the expected direction.

(2-a) Mehr als die Hälfte der Studenten hat (nicht) in der Mensa gegessen…
  more than half of the students has (not) in the mensa eaten…
  ‘More than half of the students did (not) eat in the mensa…’

(2-b) Weniger als die Hälfte der Studenten hat (nicht) in der Mensa gegessen…
  fewer than half of the students has (not) in the mensa eaten…
  ‘Fewer than half of the students did (not) eat in the mensa…’

The main eyetracking experiment tested sentences such as (2a/b) with the main verb occurring only after the negation with sentences with the verb aß (ate) in verb second position before the negation. The verb position was manipulated to investigate whether effects of quantificational complexity could show up even before the verbal predicate was
encountered. To our surprise, effects of semantic complexity only showed up at the very end of the sentences and during rereading.

To summarize, the results of two studies on the time course of scope interpretation revealed essentially non-incremental effects. The processing of quantifier scope thus seems to depend on a larger domain than the quantifier phrases themselves. Instead, quantifiers seem only to be interpreted with respect to scope after a complete sentence. If time allows, I will contrast the non-incremental processing of scope information with results from experiments showing highly incremental, predictive processing of quantificational restriction. Taken together, our experiments suggest a qualitatively different time course for interpreting the scope and the restrictor argument during online semantic processing.

REFERENCES


Language is our primary means of communication, yet most utterances listeners encounter are highly underspecified with respect to what the speaker intends to convey. Nevertheless, listeners routinely and without much apparent effort make the right kinds of inferences about the speaker’s intended meaning. How they do so is very much an unresolved question. In this talk, I will argue that it is by virtue of listeners’ capacity to efficiently make use of contextual information that communication can proceed at the remarkable rate at which it does. I will present a series of studies that investigate the role of context in the distribution, strength, and processing of a particular type of inference as a case study: scalar implicatures, i.e., the inference from “some” to “not all” as in (1).

(1)  
Anne: How was the exam?  
John: Some of the students failed.  
Inference: Not all of the students failed.

Contrary to the traditional view in linguistics, I will show converging evidence from multiple methodologies – including controlled behavioral experiments and analyses of naturally occurring speech combined with large-scale web-based survey studies – that these scalar inferences are systematically context-dependent. For example, inference strength is modulated by interlocutors’ conversational goals as well as by the alternative utterances (quantifiers, numerals) that are assumed to be available to the speaker. This context-dependence is also reflected in online processing, such that the speed with which scalar implicatures are computed is modulated by the support that the implicature receives from available contextual cues.
This talk will discuss the role of event structure and event duration in language comprehension. Several studies are reported, which investigate the contrast between state and event descriptions in sentences and short narratives. When compared to state descriptions, event sentences appear to take longer to process, consistent with the fact that events have internal causal structure and are thus expected to require more time to understand (Gennari & Poeppel, 2003). However, both states and events are also modulated by the duration of the referred eventuality and the different kinds of temporal relations that they can establish with other events in the context. For example, when presented in larger contexts (e.g., subordinate constructions), event and state sentences display different processing consequences that are consistent with their internal temporal properties. States may overlap other events in the discourse, whereas events do not tend to do so (Gennari, 2004). Similarly, when events are very short (punctual), they may require less time to process compared to long states (Coll-Florit & Gennari, 2011).

To investigate whether event structure and event duration are orthogonal to each other, we conducted several experiments contrasting different length of the same event (long vs. short). To this end, we constructed materials like (1) in which the discourse relations and the event referred stay the same and only the event duration interpretation varies due to minimal changes in the preceding context (plausibility didn’t differ across conditions).

(1) Lisa was moving to a new flat near the university.

**Long condition:** John spent his morning there.
**Short condition:** John spent an hour there.

He spent all that time assembling her bed.

We tracked participants’ eye-movements while looking at objects on the screen containing only one object related to the story being heard (e.g., bed, for example (1)). Participants only heard one condition for each item (long or short). Results indicated that first fixation durations on the relevant object (bed) were longer for the long condition while hearing her bed ($p < .05$). These results were replicated in another study using a different set of materials in which the scale of the events’ duration was longer (e.g., building a house in two weeks vs. a month).

In other studies, we used the same stimulus materials in a probe recognition task. After reading stories like (1) (with an additional final sentence to avoid recency), participants were presented with words (e.g., bed), and were instructed to indicate whether the word had occurred in the story. The word probes could either be from the critical verb phrase (*assembling her bed*) – late probes – or from the beginning of the story (e.g., flat) – early probes. In the latter case, no effect of duration should be observed if the duration change only applies the critical event, rather than the whole story. We found as expected, that participants were faster in recognizing late probes in the short-version of the story than in the long-version ($p < .05$), but no difference was found for early probes. These results indicate that duration effects prevail when the same event and narrative structure are used in the stimuli,
suggesting that the representation of an event’s internal development is more difficult for longer events.

Finally, we investigated events and states of different duration by comparing long vs. short events together with long vs. short states, as in (2). The experiment was conducted in Spanish and took advantage of the contrast between different auxiliary be-verbs (ser/estar) indicating temporary and persistent states.

(2) States:  
My son is (currently) very pale (my hijo está muy pálido).  (short state)
My son is (always) very pale (my hijo es muy pálido).  (long state)

Events:  
John wrote a note on his notebook.  (short event)
John wrote a thesis on psychology.  (long event)

These sentences were followed by a pronoun such as this, referring to the preceding sentences (e.g., This worries me). The reading times of the pronoun (referring back to the preceding state or event) indicated an effect of duration for both states and events, suggesting that duration effects are independent of event structure.

Taken together, these results suggest that effects of event structure and duration are independent of each other. We will discuss possible underlying mechanisms that may give rise to these effects, such as semantic complexity and indeterminacy.

REFERENCES


At the heart of this talk will be a set of complex Reaction Time and fMRI experiments that deployed a verification task with quantificational sentences and quantity-containing scenarios. These will be described against the background of a series of multimodal experiments in healthy adult subjects and in patients with Broca’s aphasia. I will report recent work that had two goals:

1. to study the relation between linguistic and numerical processes in the brain (anatomical localization, and the neural dynamics of verification),

2. to distinguish between semantic analyses (theoretical adjudication). This was made possible as among other things, we studied the temporal and neural dynamics of the verification of comparatives, with the hope of distinguishing between different analyses of *less*-comparatives.
The ability to determine the number of elements in some set is one of the cognitive abilities common to most humans (including infants) who also share it with other species (Antell, Keating, & Ellen, 1983; Ward, 2015; Wynn, 1992). Many natural languages allow their users to communicate number distinctions through grammatical means: formal morpho-syntactic features associated with number-related concepts like singularity and plurality (for an overview, see Corbett, 2000). The link with an extra-linguistic system of numerical cognition makes grammatical number an attractive subject for psycholinguistic research. However, to date there have been relatively few studies investigating the process of extracting number meaning from grammatical information on-line.

One such study was reported in Berent et al. (2005). They introduced a new experimental technique and used it to show that grammatically encoded number can interfere with the ability to assess the cardinality of a visually presented set of elements. In the experiment, grammatically singular and plural nouns were displayed on the screen either as a single item (e.g., “dogs”) or repeated twice (e.g., “dogs dogs”). The participants were asked to count how many items they see on the screen by pressing buttons indicating “one” or “two”. Their reaction times were measured. The data showed that when the morphological number (singular/plural) was incongruent with the perceptual number (one word/two words) the participants’ responses were significantly slower than for number-neutral control items (meaningless strings of letters, e.g., “zzzzzz”). However, these effects were obtained only for words with plural morphology. That is, when a single item was presented on the screen, the responses were on average longer if the item was a plural noun (e.g., “dogs”) than if it was a singular noun (e.g., “dog”) or a number-neutral string of letters (e.g., “zzzzzz”). The authors interpreted this Stroop-like effect as suggesting that the semantic number of nouns is extracted automatically from their morphology and that singular forms in Hebrew (the language of the experiment) are unspecified for number semantically as well as morphologically. Patson & Warren (2010) extended the technique by showing that it gives reliable results when applied to words presented in context. They applied it to investigate the interpretation of singular nouns in the scope of a distributive or collective operator. The results suggest that singular nouns in distributive contexts are actually treated as semantically plural – it took longer for the participants to decide that there was one word on the screen when the word was a part of a distributive predicate than when it was a part of a collective predicate, even though in both cases the noun was morphologically singular.

Berent et al. (2005) worked with native speakers of Hebrew and Patson & Warren’s (2010) study was done in English. In my study I applied this technique to Polish – a language typologically different from both Hebrew and English. In a series of experiments I attempted to replicate the basic effects reported previously for the two other languages and to test the usefulness of the technique as a tool for studying the contribution of different linguistic factors (morphology, lexical semantics, context) for number-related interpretation. In particular, I investigated mismatches between grammatical number and interpretation (pluralia tantum, mass nouns, group nouns), the processing of plural nouns in the scope of
negation and the potential conceptual links between verbal aspect and nominal number. I will discuss the results and conclusions of the experiments.

REFERENCES


Linguistic pathway to multiplication
Katalin É. Kiss and Tamás Zétényi
Research Institute for Linguistics of the Hungarian Academy of Sciences
Contact: ekiss@nytud.hu

GOAL Whereas it is a well-established fact that preschoolers, and even human infants can perform intuitive addition and subtraction, it is an open question whether children are capable of multiplicative operations on sets before receiving formal training (Barth et al. 2009; McCrink & Spelke 2010). What makes evidence of intuitive multiplication hard to obtain is that in the visual and auditive domains multiplication is often indistinguishable from repeated addition. We provide experimental evidence that multiplication operations are routinely processed by children prior to schooling: they are encoded by syntactic means in sentences involving distributive quantification such as the Hungarian sentences in (1)-(3). Our results also outline the acquisition path of multiplication.

BACKGROUND Hungarian has three means of expressing distributivity. The wide-scope quantifier can be associated with one of two distributive-key markers: either the universal predeterminer mind (1) or the distributive enclitic is (2). Alternatively, the distributive share can be marked by numeral reduplication (3).

(1) Mind három néni két kutyát sétáltat.
every three woman two dog-ACC walks
‘Every one of the three women is walking two dogs.’

(2) Három néni is két kutyát sétáltat.
three womanDIST two dog-ACC walks
‘Three women each are walking two dogs.’

(3) Három néni két-két kutyát sétáltat.
three woman two-two dog-ACC walks
‘Three women are walking two dogs apiece.’

The interpretation of these sentences requires multiplication. Prior to formal training and the memorization of multiplication tables, multiplication problems are solved by calculation: the child has to identify the distributive key and the distributive share, and has to add as many copies of the distributive share as specified by the distributive key. We tested whether children can calculate the meanings of sentences of type (1)-(3) prior to math education, and in what order these distributive constructions are acquired.

THE EXPERIMENT

Participants: 54 children aged 5-7, of 3 age groups: a middle group and a big kids’ group of a kindergarten, and a 1st class of an elementary school at the beginning of the schoolyear.

Materials, procedure: The child first had to judge the truth value of 6 distributive constructions, 2 of each type in (1)-(3) (and 6 fillers), coupled with pictures shown on a computer. Then the experimenter asked the child to act out 6 test sentences of type (1)-(3) (and some fillers) with toy figures. She tested the child’s interpretation of the sentence by asking how many toy figures (e.g., in the case of (1), how many dogs) she needs to set up the situation.

Results: TVJ: The passive comprehension of multiplicative sentences was nearly complete in all three age groups.
Acting-out: 25% of children could calculate distributivity/multiplication safely (with one or no mistake) at the age of 5, and 62% at the age of 7. Children acquire the distributivity-key role of mind, the lexically transparent general universal quantifier of Hungarian, before learning the distributive-key-marking function of the clitic is, and the distributive-share-marking function of numeral reduplication (Table 1).

Table 1. Correct multiplication of the distributed share in 3 types of distributive constructions

We found that children understand the procedure of multiplication (i.e., they can identify the multiplier and the multiplicand, and know that the multiplier shows the number of the copies of the multiplicand to be added up) before they are able to calculate the product mentally. 24% of the children giving correct answers solved the tasks by multiplying the set of real objects representing the multiplicand, and counting their sum. E.g., when they heard the sentence Three women each are walking two dogs, and were asked How many dogs do you need?, they set up three toy women in front of them, and asked for two toy dogs, then for two more dogs, and again for two more dogs, before announcing: I need six dogs. 30% of them multiplied sets of their fingers, instead of sets of objects. 46% of the children giving correct answers were able to calculate the result of multiplication mentally, without the help of their fingers.

The proportions of the three strategies changed with age: 70% of the first graders answering correctly used mental calculations, which was used only by 47% of the big kids and 25% of the small kids. The small kids mainly reached the solution by finger counting or by manipulating the objects (39-39%).

Table 2. Strategies of calculating the product of multiplication

THEORETICAL IMPLICATIONS The cognitive and linguistic aspects of quantification are inseparable, and have to be investigated in tandem.
REFERENCES
Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Nonsymbolic arithmetic in adults and young children. *Cognition, 98*, 199–222.
GOAL AND DATA. In the comprehension of English, processing is more costly for semantically complex than simple events (see Bott, 2010; Brennan & Pykkänen, 2010; Gennari & Poeppel, 2003; Pickering et al., 2006; Piñango et al., 1999; Piñango et al., 2006; Todorova et al., 2000a, 2000b, among others). Very little is known about the processing of events in languages, which express event structure with overt perfective/imperfective morphology. Polish as a language with a rich system of aspectual morphology is a good testing ground for investigating how the parser reacts to different patterns of interaction of lexical aspect and grammatical aspect in building temporal interpretation. In our study we present the results of a self-paced reading time and an eye-tracking experiment in which we wanted to test how the parser reacts to differences in the complexity of event structure of verbs caused by different patterns of interaction of lexical aspect and grammatical aspect in Polish. In order to test this we included the following conditions and comparisons in our experiments:

**Condition 1 (1-place imperfective verb denoting a process):**

Zosia powiedziała, że, nadąsany dzieciak szlochał
Zosia said that sulky child sobbed-IMPF
głośno w piaskownicy.
loudly in sandpit

**Condition 2 (1-place imperfective verb denoting an iterated process):**

Nina powiedziała że, nadąsany dzieciak tupał głośno
Nina said that sulky child stamped-ITER loudly
w podłogę.
in/against the floor

**Condition 3 (1-place semelfactive verb denoting a punctual event):**

Kasia powiedziała, że nadąsany dzieciak tupnął
Kasia said that sulky child stamped-SEMEL
głośno w podłogę.
loudly in/against the floor

**Condition 4 (1-place perfective verb denoting an event consisting of a process and an inception/culmination):**

Sylwia powiedziała, że nadąsany dzieciak zaszlochał
Sylwia said that sulky child sobbed-PERF
głośno w piaskownicy.
loudly in sandpit
Comparison 1: Condition 1 and Condition 2
Comparison 2: Condition 1 and Condition 4
Comparison 3: Condition 2 and Condition 3
Comparison 4: Condition 3 and Condition 4

In comparison 1, both verbs are imperfective (there is no difference in the morphological aspectual complexity) but in condition 1 the verb expresses a simple process and in condition 2 the verb describes an iterated process, which has a more complex event structure. In this comparison the manipulation of the event structure is purely lexical. In comparison 2, condition 1 the imperfective verb is morphologically and semantically less complex than its perfective counterpart in condition 4. In comparison 3, the iterative verb is more complex only at the lexical level, whereas its semelfactive counterpart is more complex at the morphological level but this morphological complexity leads to a simplification of an event structure of the verb. In comparison 4, both verbs are morphologically more complex but in condition 4 the perfective prefix increases the semantic complexity of an eventuality whereas in condition 3 the semelfactive morpheme decreases the semantic complexity of an eventuality. Based on earlier literature reports, we should expect that the conditions which are more semantically complex should be more difficult to process in all the comparisons. This prediction is not uncontroversial though, because we know very little about how the parser reacts to the morphological aspectual complexity in languages with a rich system of aspectual morphology. Moreover, we know from typological literature that in languages, which distinguish between perfective and imperfective grammatical aspect these grammes differ in terms of markedness. Marked elements are described by typologists as semantically more specified and determinate while the opposed unmarked elements are semantically indeterminate. In the aspectual markedness hierarchy, perfective aspect is placed higher than imperfective aspect and among imperfective verbs, iterative ones are more marked than noniterative ones (Battistella, 1990, p. 29). Given that unmarked elements are semantically unspecified, it is plausible to assume that they are more difficult to process than their marked counterparts, since they are more demanding for the working memory and their interpretation relies on context.

RESULTS. In the self-paced reading experiment, we obtained significant effects for Comparisons 1 and 2, where imperfective processes scored higher reading times than iterative verbs and perfective verbs. This result is particularly surprising for the second comparison, because perfective verbs were morphologically and semantically more complex than their imperfective counterparts. This result may be related to the lowest position of imperfective process in the markedness hierarchy. Being unmarked, imperfective processes are semantically underspecified and we cannot compute their meanings online so we keep them in the working memory waiting for some clues from the context to determine their ultimate interpretation. Interestingly, we obtained significant effects in reading time and regression measure in the eye-tracking experiment in the same comparisons (1 and 2), but here perfective verbs and iterative verbs received shorter reading times measures than imperfective verbs but crucially imperfective processes scored significantly more regressions out of the last Interest Area (IA) and significantly more regressions into the preverbal Interest Areas than perfective verbs and iterative verbs. Given that unlike in the self-paced reading experiments, in the eye-tracking experiments shorter reading times do not necessarily mean that a critical item is less difficult to parse, we conclude that imperfective processes in condition 1 are read quickly because they are underspecified for meaning and the parser searches for contextual specification, which is manifested in our experiment in more
regressions out of the last IA and more regressions into preverbal IA in the first condition with imperfective processes.

REFERENCES


Minimum and maximum readings in degree questions with modals

Yaron McNabb
Utrecht University
Contact: yaron.mcnabb@gmail.com

Scharten (1997) and Musolino (2004) were the first to show that when numerals are embedded under universal modals like *have to* in (1a), they receive a lower-bounded, “at least” reading, and when they occur with existential modals like *be allowed* in (1b), they receive an upper-bounded, “at most” reading.

(1) a. You have to drive 50kph on this motorway.
   b. You’re allowed to drive 120kph on this motorway.

While it’s standard to derive both lower- and -upper bound readings, at least for universal modals, via different scope configurations (see Kennedy (2015) for an implementation), Abrusan and Spector (2011) point out that there’s a parallel between the readings in (1) and degree questions with modals. In a context where a person is being tailgated because he’s driving too slow, asking the question in (2a) is odd with the existential modal *be allowed*. The opposite is true in a context in which a person is stopped by the police officer for driving too fast; here *be allowed* is well-formed while the universal *have to* is odd. Abrusan and Spector (2011) further argue that these modals contrast with a third set of modals, like *should*, which are in turn compatible in both contexts, as shown in (2).

(2) a. Why is the person behind me tailgating me? How fast {#do I have to / should I / am I allowed to} drive?
   b. Officer, why did you stop me? How fast {#do I have to / should I / am I allowed to} drive?

The aim of this study is to investigate a larger number of modals that belong to the three groups Abrusan and Spector (2011) classified. 110 native English speakers, recruited via Amazon’s Mechanical Turk, read short scenarios like the one in (3), which varied in their bias toward a lower-bound (MINIMUM) or upper-bound (MAXIMUM) reading, and were asked to evaluate how coherent was the question given the context on a Likert scale of 1 (completely incoherent) to 7 (completely coherent).

(3) a. Anna suffers from hay fever every year. This year, when the symptoms started, she went to her doctor and told him that she took 3 antihistamines every day. The doctor prescribed her more pills. When Anna saw the prescription, she asked him: *Oh, How many antihistamine pills do I need to take per day?*  
   MINIMUM

   b. Anna suffers from hay fever every year. This year, when the symptoms started, she went to her doctor and told him that she took 5 antihistamines every day. The doctor prescribed her fewer pills. When Anna saw the prescription, she asked him: *Oh, so how many antihistamine pills do I need to take per day?*  
   MAXIMUM

Nine versions were created for each type of context, each followed by a question with a different modal, three of each belonging to one of the three hypothesized groups: (1)
EXISTENTIALS (max-biased): allowed to, can, may; (2) STRONG UNIVERSALS (min-biased): required to, need to, have to; and (3) WEAK UNIVERSALS (neutral): supposed to, should, must. Each participant read 18 target items, 20 control items, and 45 filler items.

The data were analyzed with mixed-effects ordered prohibit regression models with random effects for subjects and items. The box plots below shows that degree questions with strong universal modals were more coherent in upper-bound contexts than in lower-bound contexts, that degree questions with existential modal were more coherent in lower-bound contexts than in upper-bound context ($p < 0.001$ for all comparisons), and that context only marginally significantly ($p = 0.02$) affected the coherence of weak universals in degree questions.

![Box plots showing coherence of modals in different contexts](image)

This study provides experimental evidence for a three-way split between root modals with respect to the preferred readings they give rise to: minimum, maximum or either. This pattern in terms of a lower- or upper-bound bias is parallel to the preferred readings of these modals with combination with other degree constructions, such as comparatives (Heim, 2006), as well as their scopal preferences under negation and their licensing in downward entailing environments (Iatridou & Zeijlstra, 2013).

REFERENCES


‘Many’ expectations about ‘few’ balls

Anthea Schöller and Michael Franke

University of Tübingen

Contact: anthea.schoeller@uni-tuebingen.de; mchfranke@gmail.com

ABSTRACT Work in psychology suggests that the use of context-dependent expressions like few and many is dependent on speakers’ prior expectations. In real-world contexts, however, it is not possible to manipulate expectations in a controlled manner. We move to more abstract contexts and present urns of varying content. We show that prior expectations can be manipulated in this way and that they are a significant predictor of the acceptability of few and many.

LANDSCAPE One key property of vague quantifiers like few and many is that their denotation can vary extremely across contexts. The number of spectators described by many in (1a) will be much larger in a conversation about the final of the European Championship than for a match of the local football club. The same variability can be found in few in (1b) as well.

(1) a. Many people watched the football match on Sunday.
   b. Few of the guests at the party drank beer.

Literature in psychology informally suggests that world knowledge influences the use of vague quantifiers or gradable adjectives (Clark, 1991; Moxey & Sanford, 1993). These insights have not been incorporated in formal semantics yet. Especially theories based on a degree semantics like Hackl (2000), Kennedy (2007), and Solt (2009) and also Partee (1989) formulate truth conditions in terms of whether a cardinality or degree exceeds a contextually given threshold. How exactly this threshold is influenced by the context is not spelled out. We set out to identify one aspect of the context to which we ascribe a major role in the production and interpretation of these vague expressions.

HYPOTHESIS Following Clark (1991), Fernando and Kamp (1996), and Frank and Goodman (2012) we assume that the context is integrated into the reasoning about language use in the form of prior expectations. The way we use vague quantifiers like few and many is influenced by the quantities we have expected and consider to be normal in the respective situation.

EXPERIMENTAL INVESTIGATION Since each speaker has individual experiences of a real-world context, it is very hard if not impossible to manipulate expectations in a controlled manner. This is why we move to using a clearly manipulable fixed contextual prior setting. A sample of the visual material can be found in Figure 1. A character draws 10 red and blue balls from an urn. The draw’s outcome is displayed visually. The character describes the NUMBER of blue balls drawn [1-10] with a statement including the QUANTIFIER few or many as in (2). We explicitly mention the content of the urn in a for-phrase to make the prior salient.

(2) For a draw from an urn with that content, [few|many] of the balls I drew are blue.

To manipulate PRIOR EXPECTATIONS, we present a picture of the urn from which the balls were drawn. The urn’s content varies. From a total of a 100 balls either 25, 50, 75 or 90 balls are blue, the rest red. We recruited 90 participants via Mechanical Turk and asked them to rate on a 7 point scale whether the quantified sentence is a good description of the situation.
**Figure 1.** Sample pictures of the drawn balls and an urn with 25, 50 and 75 blue balls

**PREDICTIONS** We expect that NUMBER is a significant predictor of participants’ behaviour. *Few* is applicable to low numbers whereas *many* will be rated highest for numbers in the upper range. We expect that PRIOR EXPECTATIONS have a significant effect since both quantifiers express, or can express, that a cardinality is lower or higher than expected. When prior expectations are low [25 blue balls in urn], both quantifiers are applicable to lower numbers than when prior expectations are high [75 or 90 blue balls].

**RESULTS** Mean ratings of each PRIOR-QUANTIFIER pair can be found in Figure 2. For each of the quantifiers *few* and *many* we specified a mixed linear effects regression model predicting ratings which included the main effects NUMBER OF BLUE BALLS (1-5 for *few*, 4-9 for *many*) and PRIOR EXPECTATIONS (25, 50, 75 or 90 blue balls out of 100).

*FEW* For *few*, the model included the fixed effect NUMBER. Participants gave significantly lower ratings for a rising number of balls ($\beta = -0.52, SE = 0.08540, p < 0.001$). The factor PRIOR EXPECTATIONS was significant, too. Higher prior expectations led to higher ratings ($\beta = 0.33, SE = 0.089, p < 0.002$). This effect was modulated by a significant interaction between NUMBER and PRIOR EXPECTATION ($\beta = 0.18, SE = 0.077, p < 0.02$). A higher expected cardinality seems to have the effect that *few* is applicable to higher numbers.

*MANY* Ratings are significantly higher for an increasing NUMBER of balls ($\beta = 0.72, SE = 0.063, p < 0.001$). We found a significant effect of the factor PRIOR EXPECTATIONS as well. Higher prior expectations lead to higher ratings ($\beta = 0.46, SE = 0.087, p < 0.001$). However, there is no significant interaction between NUMBER and PRIOR EXPECTATION.

**DISCUSSION** We successfully manipulated expectations in an abstract context and found that they influence the ratings of sentences including *few* and *many*. For both *few* and *many* the factor PRIOR EXPECTATIONS has a significant effect which confirms that both express that a cardinality is lower or higher than expected. However, only for *few* we find a significant interaction of NUMBER and PRIOR EXPECTATIONS. When a low cardinality is expected, the ratings decrease more for a rising number of blue balls than when a higher cardinality is expected. We conclude that the range *few* applies to is more restricted than *many*’s. *Few*’s range can be slightly shifted, but in general *few* only applies to small cardinalities, whereas *many*’s range is more flexible. Overall, we successfully applied a new experimental paradigm and confirmed that contextual information can be integrated in the form of prior expectations. However, this insight is only a first step and more work is necessary. To substantiate our claim we plan to extend the investigation to further constructions like *few* and *many* in combination with *compared to* frame setters. Other open questions are how to integrate prior expectations into the formal semantics and how they interact with the threshold proposed by degree semantic theories.
Figure 2. Results of the rating study: mean ratings of each PRIOR-QUANTIFIER pair

REFERENCES


This study investigates how quantifying expressions (numbers vs. some) are interpreted in Turkish. We address two competing accounts regarding the interpretation of the quantifying expressions. According to the lower-bounded account, both numbers and scalar terms are initially interpreted with their lower-bounded semantics (i.e., I ate three apples means I ate at least three apples; I ate some apples is compatible with a context where the speaker has eaten all of the apples), and the exact meaning interpretation arises through pragmatic inferences (Horn, 1972; Gadzar, 1979). The exact-semantics account suggests that numbers are unlike other scalar terms, as they lead to the exact interpretation by default (Horn, 1992). In line with this account, Huang, Spelke, and Snedeker (2013) showed that numbers are interpreted with their exact meaning whereas quantifiers are interpreted with their lower bounded meaning. We report the first experiment of an ongoing project that focuses on the interpretation of numbers and three different scalar terms for some in Turkish (i.e., elmalarin birazibir kismi/bazlari) with a task modeled on the covered-box task in Huang et al., (2013). The aim in testing these three different terms was to find out whether our intuition is correct that each of them would lead to different interpretations. Here we focus on the term birazi that is ambiguous between some of and a few/a little of, which might rule out the upper-bounded interpretation.

We presented 20 Turkish speakers with a scene with three cards and a sentence on the screen beneath the scene, and asked them to click on the correct card depicting the sentence they read. Of the three cards, one was covered while the content of the other two were visible to the participants. Each card depicted two characters and a set of identical objects. In some cases, all of these objects belonged to one of the characters; in others, they were split between the characters. Similar to Huang et al. (2013) study, there were three conditions for the visual scenes. In NONE-SOME trials, participants were presented with one subset match with a picture where both figures had some but not all of the objects, one empty set match where the target figure has none of the objects and the other figure has all of the objects, and a third covered box. In SOME-ALL trials, participants were presented with one subset match with a picture where both figures had some but not all of the objects, one total set match where the target figure has all of the objects and the other figure has none of the objects, and a third covered box. In NONE-ALL trials, participants were presented with one total set match where the target figure has all of the objects, one empty set match where the target figure has none of the objects and the other figure has all of the objects, and a third covered box. As for the test sentences, we had two conditions, the scalar condition with some (birazi) and the number condition, as in (1) and (2) respectively.

(1) Hangi resim-de elma-larin biraz-ı Ayşe’de?
which picture-Loc apple-Pl-Gen some-Poss.3sg Ayse-Loc
‘In which picture Ayse has some of the apples?’

(2) Hangi resim-de elma-larin üç-ü Ayşe’de?
which picture-Loc apple-Pl-Gen three-Poss.3sg Ayse-Loc
‘In which picture Ayse has three of the apples?’
In this preliminary study we found that number words were interpreted with their exact semantics (Figure 1), which is consistent with the findings in English (Huang et al., 2013). However, contrary to these findings, the scalar term *some/biraz* led to an exact interpretation (Figure 2). This might be due to the ambiguity of the term *biraz*. This term might have been interpreted as *a few of*, in which case the card depicting one of the characters owning all of the objects would be a mismatch. Currently, we are testing the two other terms, namely *bir kismi* and *bazlari* to see whether they would lead to a lower-bounded interpretation, just like *some* in English.

**Figure 1. Numbers**

![Graph 1]

**Figure 2. Quantifiers**

![Graph 2]

REFERENCES


The notional category ‘event’ is important for research across the social and behavioral sciences: linguists say that sentences are about events, philosophers debate the metaphysics of event identity, cognitive psychologists posit event concepts to explain how creatures like us represent and reason about the world, and developmental psychologists ask how we come to have those concepts. I present work that is part of a project investigating the relationship between compositional event semantics and event representations in the psychologist’s sense. I begin by considering semantic parallels between the nominal and verbal domains: just as we can distinguish reference to substance, atomic objects, and pluralities in the nominal domain (Parsons, 1979), we can distinguish reference to process/activity, atomic events, and pluralities in the verbal domain (Bach, 1986). Recently, these ontological parallels have been used to help explain how quantification with cross-categorial ‘more’ works (Wellwood, 2015). In this talk, I link atomicity to a principle of ‘unity and organization’ characteristic of object representation (see Rips & Hespos, 2015 for a review), and extend this to event representation; the basic idea is that representing a thing as an object or an event is to represent it as something worth counting (cf. Koslicki, 1997). I present the results of experiments involving static and dynamic displays with arbitrary or non-arbitrary spatial (static) and temporal (dynamic) gaps (cf. Maguire et al., 2011), and discuss how these are mapped to mass/count language with novel nouns (see also Prasada et al., 2002, Barner & Snedeker, 2006), as well as to language introducing similar semantic distinctions with novel verbs (gleeb every second or so / gorp around a little).

REFERENCES
Many of the natural language expressions have contextually variable interpretations. For instance, all the examples in (1) can convey different information in different contexts about how the predicate is interpreted as a particular value.

(1)  
   a. John is *tall*.
   b. This room is *full*.
   c. There are *100* people on the waiting list.

The interpretation of relative adjectives (e.g., *tall* in 1a) depends on a contextually determined comparison class. Absolute adjectives (e.g., *full* in 1b) and numerals (e.g., *100* in 1c) can have either precise or approximate interpretations depending on the context.

In this talk we present experimental data to explore two questions. First, we ask whether the seemingly identical contextual variability arises from different sources. The empirical focus of this investigation is the comparison between the relative and absolute adjectives (e.g., 1a and b). Adapting the visual world eyetracking paradigm from Sedivy et al. (1999), our experiment investigated the interaction of context and adjective meaning in a reference resolution task involving definite descriptions with relative and absolute adjectives used as restrictive modifiers. Previous studies using this paradigm showed that a *Contrast* object in the visual context (e.g., a short glass) facilitated the processing of a relative adjective modifier (e.g., when participants were asked to “point to a tall glass”). Building upon this previous finding, our studies showed that relative and absolute adjectives show different sensitivity to the presence of the contrast object. In particular, the contrast object facilitates absolute adjective processing only when the target satisfies the adjective description imprecisely. We explore the hypothesis that the contextually variable interpretations for relative and absolute adjectives arise from different sources.

The second question we investigate is the processing cost involved in deriving precise vs. imprecise interpretations. The results from the first set of studies on absolute adjectives provide some evidence to show that participants do not privilege the precise interpretation of absolute adjectives over the imprecise interpretation. It has also been proposed informally before that imprecise interpretations are cognitively less costly (Krifka, 2009), but there was no empirical data to support this claim. In the second set of studies, we focus on the interpretation of numerals. Contexts were created that were compatible with both precise and imprecise interpretations of a numeral. After viewing the context, participants self-paced-read a target sentence that contained a numeral. Adverbs like *precisely* and *approximately* were used to in the target sentence to trigger a precise or imprecise interpretation. Reading time data showed that numerals following the adverb *approximately* had much longer RT than the same numerals following the adverb *precisely*, suggesting that imprecisely interpreting a numeral is actually more costly, contrary to previous claims. We discuss this finding in terms of the general processing mechanism that sets up the range of possible values of imprecise interpretations.
REFERENCES
