Investigating scalar implicatures in a truth-value judgment task: Evidence from event-related brain potentials:

Supplementary Materials

Maria Spychalska, Jarmo Kontinen, Markus Werning

Language, Cognition, Neuroscience

1 An example graphical visualization of the cluster-based permutation tests results

Figures (1) and (2) graphically present the results of the cluster-based permutation test for the comparison *Some-Infelicitous* vs. *Some-True* with marked significant clusters and time-windows.

Figure 1: The significant negative clusters (marked red) for the comparison *SI* vs. *ST* determined by the permutation test for (a) “pragmatist” and (b) “logicians”.

(a)

(b)

Figure 2: The significant positive clusters (marked red) for the comparison *SI vs. ST* determined by the permutation test for (a) “pragmatist” and (b) “logicians”.

2 The factorial analysis in the pre-defined time-windows

To compare the results of the cluster-based permutation tests with the more traditional approach to the statistical analysis of the EEG data we additionally performed a repeated measures analysis of variance. For this purpose we computed the mean amplitude over the whole scalp¹ in a pre-defined time-window of 250 – 450 *ms* for the N400 effect and 550 – 950 *ms* for the PNP effect. These time-windows were chosen based on the literature, the inspection of the grand averages as well as the results of the permutation testes. All our dependent variables were normally distributed. Greenhous-Geisser was applied whenever the assumption of sphericity was not met and the p-values of all pairwise comparisons were Bonferroni corrected. The Levene’s tests for each of the repeated measures variables were not significant, thus the variances are assumed to be homogeneous for all the compared levels.

N400 time window 250 – 450 *ms*

In the time-window of 250 – 450 *ms*, the all-factorial analysis (see Table 1) revealed a significant effect of *Truth* ($p < .001$), but not of *Quantifier* ($p = .828$). The following interactions were significant: *Truth* $\times$ *Group* ($p = .030$), *Quantifier* $\times$ *Truth* ($p = .022$) and *Quantifier* $\times$ *Truth* $\times$ *Group* ($p < .001$). There was no significant between group difference ($p = .077$), and *Quantifier* $\times$ *Group* interaction was not significant ($p = .244$).

¹Both for the N400 and PNP effects the EOG electrodes, linked mastoids as well as the anterior electrodes: Fp1, Fp2, AF7, AF8 were excluded.

Table 1: Full-factorial ANOVA of the scalp-averaged ERPs in the time-windows of 250 – 450 *ms* and 550 – 950 *ms* post-stimulus onset. The significant *p*-values are bolded.

	250 – 450 <i>ms</i>			550 – 950 <i>ms</i>		
	F-stat	p-value	effect size	F-stat	p-value	effect size
<i>Truth</i>	$F(1.69, 88.14) = 112.964$	$p < .001$	$\eta^2 = .685$	$F(2, 104) = 38.489$	$p < .001$	$\eta^2 = .425$
<i>Quantifier</i>	$F(1, 52) = .048$	$p = .828$	$\eta^2 = .001$	$F(1, 52) = 9.334$	$p = .004$	$\eta^2 = .152$
<i>Quantifier</i> $\times$ <i>Truth</i>	$F(2, 104) = 3.964$	$p = .022$	$\eta^2 = .071$	$F(2, 104) = 4.478$	$p = .014$	$\eta^2 = .079$
<i>Group</i>	$F(1, 52) = 3.257$	$p = .077$	$\eta^2 = .059$	$F(1, 52) = 1.720$	$p = .195$	$\eta^2 = .032$
<i>Truth</i> $\times$ <i>Group</i>	$F(1.69, 88.14) = 3.913$	$p = .030$	$\eta^2 = .070$	$F(2, 104) = 4.300$	$p = .016$	$\eta^2 = .076$
<i>Quantifier</i> $\times$ <i>Group</i>	$F(1, 52) = 1.388$	$p = .244$	$\eta^2 = .026$	$F(1, 52) = .366$	$p = .548$	$\eta^2 = .007$
<i>Quantifier</i> $\times$ <i>Truth</i> $\times$ <i>Group</i>	$F(2, 104) = 12.294$	$p < .001$	$\eta^2 = .191$	$F(2, 104) = 2.099$	$p = .128$	$\eta^2 = .039$

Table 2: The repeated measures ANOVA for the scalp-averaged ERPs in the pre-defined time-windows and for each of the quantifiers. The significant *p*-values are bolded.

		<i>some</i>			<i>all</i>		
		F-stat	p-value	effect size	F-stat	p-value	effect size
250 – 450	<i>Truth</i>	$F(1.79, 92.91) = 60.634$	$p < .001$	$\eta^2 = .538$	$F(2, 104) = 72.577$	$p < .001$	$\eta^2 = .583$
	<i>Truth</i> $\times$ <i>Group</i>	$F(1.79, 92.91) = 15.039$	$p < .001$	$\eta^2 = .224$	$F(2, 104) = .764$	$p = .469$	$\eta^2 = .014$
	<i>Group</i>	$F(1, 52) = 3.689$	$p = .060$	$\eta^2 = .066$	$F(1, 52) = 2.556$	$p = .116$	$\eta^2 = .047$
550 – 950	<i>Truth</i>	$F(2, 104) = 21.669$	$p < .001$	$\eta^2 = .294$	$F(2, 104) = 28.807$	$p < .001$	$\eta^2 = .356$
	<i>Truth</i> $\times$ <i>Group</i>	$F(2, 104) = 6.081$	$p = .003$	$\eta^2 = .105$	$F(2, 104) = .738$	$p = .481$	$\eta^2 = .014$
	<i>Group</i>	$F(1, 52) = 2.069$	$p = .156$	$\eta^2 = .038$	$F(1, 52) = 1.214$	$p = .276$	$\eta^2 = .023$

In order to investigate the interactions we analyzed each quantifier separately. For both quantifiers *Truth* was a significant factor ($p < .001$ in each case), but *Group* was not significant, and the *Truth* $\times$ *Group* interaction was only significant for the quantifier *some* ($p < .001$) (Table 2). All pairwise comparisons of the levels of *Truth* were significant for *some* ($p < .001$ for each comparison), and for *all* ($p \leq .001$). The analysis of repeated contrasts revealed that the difference in the mean amplitude between the conditions *Some-True* and *Some-Infelicitous* was significantly larger for “pragmatists” than for “logicians” ($F(1, 52) = 24.278$, $p < .001$, $\eta^2 = .318$). Furthermore, the difference in the mean amplitude between conditions *Some-Infelicitous* and *Some-False* was significantly larger for “logicians” than for “pragmatists” ($F(1, 52) = 27.116$, $p < .001$, $\eta^2 = .343$). This confirms the hypothesis that whereas the *Some-Infelicitous* vs. *Some-True* N400 effect was significantly larger for “pragmatists”, the *Some-False* vs. *Some-Infelicitous* N400 effect was significantly larger for “logicians”. However, for the quantifier *all* the differences in the mean amplitude between all three evaluation conditions were independent of the group.

P600 time window: 550 – 950 *ms*

In the time-window of 550 – 950 *ms*, *Truth* ($p < .001$) was a significant factor as well as *Quantifier* ($p = .004$, with the mean amplitudes for *some* being slightly more positive than for *all*: $.359\mu V$ vs. $.092\mu V$). There was a significant *Truth* $\times$ *Group* interaction ($p = .016$) and *Quantifier* $\times$ *Truth* interaction ($p = .014$). *Group* was not significant ($p = .195$) and neither was the *Quantifier* $\times$ *Group* interaction ($p = .548$) nor the *Quantifier* $\times$ *Truth* $\times$

Group interaction ($p = .128$).

The follow-up analysis of the average ERPs for each quantifier revealed that *Truth* was a significant factor for both *some* ($p < .001$) and *all* ($p < .001$), but *Group* was not significant, and the *Truth* $\times$ *Group* interaction was significant only for *some* ($p = .003$) (Table 2). In the case of *some*, two pairwise comparisons of the levels of *Truth* were significant: *Some-False* vs. *Some-True* ($p < .001$), *Some-False* vs. *Some-Infelicitous* ($p < .001$), and the comparison between *Some-Infelicitous* and *Some-True* was marginally significant ($p = .059$). The analysis of repeated contrasts showed that the difference in the mean amplitude between conditions *Some-True* and *Some-Infelicitous* was significantly larger for “pragmatists” than for “logicians” ($F(1, 52) = 11.416$, $p = .001$, $\eta^2 = .180$). The contrast between conditions *Some-False* and *Some-Infelicitous* when comparing the two groups was marginally significant ($F = 3.926$, $p = .053$, $\eta^2 = .070$).

For the quantifier *all* both false conditions showed an increased positivity compared to the *All-True* condition, which was independent of the group: $p < .001$ for each comparison *All-False-Primed* vs. *All-True* and *All-False-Non-primed* vs. *All-True*, but there was no significant difference between the two false conditions *All-False-Primed* vs. *All-False-Non-primed* ($p = .833$).

Figure 3: PRAGMATISTS: The topographical distribution of the N400 (250 – 450 ms) and PNP (550 – 950 ms) effects in the comparisons of *S*-conditions.

Figure 4: LOGICIANS: The topographical distribution of the N400 (250 – 450 *ms*) and PNP (550 – 950 *ms*) effects in the comparisons of *S*-conditions.

Figure 5: PRAGMATISTS: The topographical distribution of the N400 (250 – 450 *ms*) and PNP (550 – 950 *ms*) effects in the comparisons of *A*-conditions.

Figure 6: LOGICIANS: The topographical distribution of the N400 (250 – 450 *ms*) and PNP (550 – 950 *ms*) effects in the comparisons of *A*-conditions.

Figures (3), (4), (5) and (6) show the topographical maps of the effects in the comparisons of *S*-conditions and *A*-conditions for both groups in the time-windows of the ANOVA.

3 The whole group analysis

Figure (8) presents grand averages for the whole tested population in each condition at nine centro-parietal electrodes. Figures (7) and (9) present topographical maps of the effects in the comparisons of the *S*-condition and *A*-conditions.

Figure 7: ALL SUBJECTS: The topographical distribution of the N400 (250 – 450 *ms*) and PNP (550 – 950 *ms*) effects in the comparisons of *S*-conditions.

Figure 8: Grand averages for all evaluation-conditions and for all subjects, shown at nine centro-parietal electrode locations, for (a) *some* and (b) *all*.

(a)

(b)

Figure 9: ALL SUBJECTS: The topographical distribution of the N400 (250 – 450 *ms*) and PNP (550 – 950 *ms*) effects in the comparisons of *A*-conditions.

4 Materials

In Table (3) we present the list of noun triples that were used to generate the stimuli. For each noun we provide its English translation and the log-frequency value based on the <http://wortschatz.uni-leipzig.de/> database, 2011. Table (4) presents the noun triples used in the filler trials, whereas Table (5) lists all the types of the filler trials that were used in the experiment. Each filler triple was used three times to generate the scenarios, but each word was used only once as a critical word. We balanced out per quantifier the number of true and pragmatically felicitous, true and pragmatically infelicitous and false filler trials. The true trials with bare numerals were varied with respect to the numeral’s interpretation (the so-called *exactly* and *at least* readings of numerals). Note that there were no *at least* trials for the numeral *five*, as five was the maximal number of objects for a given category that could be presented. The true trials with *die meisten* (*most*) included those that were consistent (felicitous trials) and inconsistent (infelicitous trials) with the implicature *not all*. Additionally, the false trials with *die meisten* or with bare numerals varied with respect to whether the critical word denoted one of the presented objects (primed) or not (not primed).

Table 3: Noun triples used in the experimental trials

<i>Noun 1</i>	<i>Freq.</i>	<i>Noun 2</i>	<i>Freq.</i>	<i>Noun 3</i>	<i>Freq.</i>	<i>Trans. noun 1</i>	<i>Trans. noun 1</i>	<i>Trans. noun 1</i>
Bänke	13	Kühe	13	Gänse	13	benches	cows	geese
Äpfel	13	Igel	14	Becher	13	apples	hedgehog	mug
Enten	12	Affen	12	Kleider	12	ducks	monkeys	dresses
Monde	13	Zebbras	13	Robben	13	moons	zebras	seals
Mäuse	13	Socken	13	Treppen	13	mice	socks	strais
Betten	12	Tiger	11	Ratten	12	beds	tigers	rats
Rehe	15	Leitern	14	Möwen	15	deers	ladders	seagulls
Bären	11	Zelte	12	Reifen	11	bears	tents	tires
Möhren	15	Ziegen	14	Blusen	15	carrots	goats	blouses
Frösche	14	Schnecken	14	Rüben	15	frogs	snails	beetroots
Schlangen	12	Fische	12	Fliegen	12	snakes	fish	flies
Hunde	11	Knochen	11	Adler	11	dogs	bones	eagels
Hasen	13	Gurken	14	Spinnen	14	hares	cucumbers	spiders
Birnen	15	Torten	15	Schüsseln	15	pears	(birthday)cakes	bowls
Trauben	14	Kekse	14	Falken	14	grapes	biscuits	falcons
Schwäne	13	Eulen	15	Lilien	14	swans	owls	lilies
Sessel	13	Geigen	15	Brote	15	armchairs	violins	breads
Mäntel	15	Hüte	14	Wespen	14	coats	hats	wasps
Schleifen	14	Bürsten	16	Krähen	15	ribbons	brushes	crows
Körbe	15	Schirme	15	Tulpen	15	baskets	umbrellas	tulips
Ballons	15	Teddys	16	Erbsen	15	balloons	teddy-bears	peas
Häuser	9	Bäume	10	Züge	10	houses	trees	trains
Engel	13	Glocken	13	Stiefel	13	angels	bells	boots
Vögel	10	Blumen	11	Türen	10	birds	flowers	doors
Federn	12	Kerzen	12	Käfer	12	feathers	candles	beetles
Bücher	10	Uhren	12	Haken	12	books	clocks	hooks
Mädchen	8	Spiegel	9	Fenster	9	girls	mirrors	windows
Schweine	13	Sterne	12	Kröten	14	pigs	stars	toads
Hefte	14	Stifte	15	Hirsche	15	copybooks	pencils	deers (males)
Koffer	11	Schuhe	11	Briefe	11	suitcases	shoes	letters
Taschen	12	Gürtel	13	Wölfe	12	handbags	belts	woolfs
Hühner	11	Eimer	13	Wurzeln	11	chickens	buckets	roots
Lampen	13	Kissen	13	Füchse	13	lamps	pillows	foxes
Puppen	13	Drachen	13	Kugeln	13	dolls	kites	glassballs
Trommeln	13	Gitarren	14	Sofas	14	drums	guitars	sofas
Tassen	14	Löffel	14	Schinken	14	cups	spoons	hams
Hämmer	17	Schwämme	17	Kiwis	17	hammers	sponges	kiwis
Gabeln	17	Kannen	17	Seifen	17	forks	jugs	soaps
Pilze	12	Blätter	12	Haie	12	mushrooms	leaves	sharks
Pinsel	14	Scheren	15	Wiegen	15	paintbrushes	scissors	cradles
Besen	14	Töpfe	14	Perlen	14	brooms	pots	pearls
Autos	8	Jungen	8	Wege	9	cars	boys	roads
Vasen	15	Muscheln	14	Zirkel	14	vases	shells	calipers
Nadeln	14	Knöpfe	14	Beeren	14	needles	buttons	berries
Schlüssel	11	Münzen	13	Steine	11	keys	coins	stones
Eier	11	Messer	11	Fahnen	11	eggs	knives	flags
Schals	14	Flöten	16	Broschen	17	scarfs	flutes	brooches
Brillen	14	Mützen	14	Bohnen	14	glasses	caps	beans
Angeln	13	Boote	12	Beutel	13	rods	boats	bags
Hosen	12	Ruder	12	Brunnen	12	trousers	paddles	springs
Fäden	12	Gläser	13	Kisten	13	threads	wineglasses	boxes
Babys	11	Flaschen	11	Brücken	11	babies	bottles	bridges
Teller	12	Hemden	13	Roller	12	plates	shirts	motorbikes
Sonnen	15	Zäune	14	Blitze	14	suns	fences	lightenings
Schafe	12	Teufel	11	Kuchen	12	sheep	devils	cakes
Schränke	14	Schaukeln	15	Elche	15	wardrobes	shovels	mooses
Bögen	14	Röcke	14	Nelken	15	bows	skirts	carnations
Jacken	13	Stühle	12	Körner	13	jackets	chairs	grains
Tische	12	Löwen	11	Decken	11	tables	lions	blankets
Pfannen	16	Sägen	15	Raben	15	pans	saws	raven
Nägel	14	Ketten	13	Tauben	13	nails	chains	pigeons
Rosen	12	Ringe	12	Bienen	13	roses	rings	bees
Schwerter	15	Helme	15	Strümpfe	15	swords	helmets	stockings
Kronen	12	Pferde	11	Schiffe	11	crowns	horses	ships
Käse	12	Kirschen	14	Hocker	13	cheese	cherries	stools
Herzen	10	Wolken	11	Netze	12	hearts	clouds	nets
Pfeile	15	Kreuze	14	Leinen	14	arrows	crosses	ropes
Esel	14	Zwiebeln	14	Öfen	15	donkeys	onions	ovens
Anker	13	Würfel	13	Segel	13	anchors	dice	sails
Kämme	17	Pflaumen	16	Zöpfe	15	combs	plumps	braids
Pfeifen	14	Masken	13	Tannen	15	pipes	masks	firs
Hörner	14	Fässer	15	Würste	15	horns	barrels	sausages
Stiere	16	Waffeln	16	Kanus	16	bulls	waffles	canoes
Raupen	15	Kegel	15	Hamster	15	caterpillars	pins	hamsters
Bonbons	14	Pflaster	13	Nüsse	15	candies	plasters	nuts
Spritzen	14	Dosen	12	Samen	13	syringes	cans	seeds
Wale	13	Palmen	13	Kreisel	14	whales	palms	spinning top
Stempel	13	Würmer	14	Pillen	13	stamps	worms	pills
Störche	15	Schlitten	13	Krebse	15	ostriches	sledges	crayfish
Biber	14	Ampeln	13	Krücken	14	beaver	traffic lights	crutches

Table 4: Noun triples used in the fillers

<i>Pair number</i>	<i>Noun 1</i>	<i>Freq.</i>	<i>Noun 2</i>	<i>Freq.</i>	<i>Noun 3</i>	<i>Freq.</i>	<i>Trans1</i>	<i>Transl2</i>	<i>Transl3</i>
1	Elefanten	12	Fahrräder	13	Skulpturen	12	elephants	bikes	sculptures
2	Handtücher	15	Hufeisen	16	Pfirsiche	17	towels	horseshoe	peaches
3	Krawatten	14	Kamine	16	Trompeten	15	ties	chimneys	trumpets
4	Pullover	13	Rucksäcke	15	Matratzen	14	sweater	backpacks	mattresses
5	Schmetterlinge	14	Delphine	15	Traktoren	14	butterflies	dolphins	tractors
6	Flugzeuge	10	Handys	10	Fotos	9	planes	mobiles	photos
7	Tropfen	12	Noten	11	Medaillen	11	drops	notes	medals
8	Teppiche	14	Regale	13	Sirenen	14	carpets	shelves	sirens
9	Bananen	13	Tomaten	13	Büsche	14	bananas	tomatoes	bushes
10	Kamele	15	Pandas	16	Schlafsäcke	16	camels	pandas	sleeping bags
11	Schildkröten	15	Zitronen	15	Klinken	15	turtles	lemons	handles
12	Narzissen	16	Ameisen	14	Streichhölzer	16	daffodils	ants	matches
13	Telefone	13	Duschen	13	Nester	15	telephones	showers	nests
14	Kakteen	16	Eichhörnchen	14	Eidechsen	16	cacti	squirrels	lizards
15	Kartoffeln	12	Zigaretten	11	Türme	13	potatoes	cigarettes	towers
16	Giraffen	15	Klaviere	16	Melonen	16	giraffes	pianos	melons
17	Murmeln	16	Wimpel	15	Hecken	14	marbles	pennant	hedges
18	Orangen	15	Strohhalme	17	Magnete	17	oranges	straws	magnets
19	Kastanien	15	Krokodile	15	Perücken	15	chestnuts	crocodiles	wigs
20	Käfige	15	Pinguine	13	Schubladen	14	cages	penguins	drawers

Table 5: Filler trials

<i>Quantifier</i>	<i>Evaluation</i>	<i>Number of examples</i>
keine	true	12
keine	false	12
die meisten	true felicitous	3
die meisten	true infelicitous	3
die meisten	false primed	3
die meisten	false not primed	3
zwei	true exactly	2
zwei	true at least	2
zwei	false primed	1
zwei	false not primed	1
drei	true exactly	2
drei	true at least	2
drei	false primed	1
drei	false not primed	1
vier	true exactly	2
vier	true at least	2
vier	false primed	1
vier	false not primed	1
fünf	true exactly	3
fünf	false primed	1
fünf	false not primed	2

In Tables (6) and (7) we report the mean log-frequencies and lengths over the words that were used in the respective conditions after the random assignment of the nouns to conditions for each subject.

Table 6: The words' mean log-frequencies for each condition after the random assignment of words

	ST	SI	SF	AT	AFP	AFN
<i>Logicians</i>	13.26	13.22	13.40	13.15	13.16	13.31
<i>Pragmatists</i>	13.18	13.19	13.35	13.18	13.18	13.37
<i>All subjects</i>	13.21	13.21	13.37	13.16	13.17	13.34

Table 7: The words' mean lengths for each condition after the random assignment of words

	ST	SI	SF	AT	AFP	AFN
<i>Logicians</i>	6.03	6.08	5.95	6.04	6.02	5.95
<i>Pragmatists</i>	5.99	6.06	5.96	6.05	6.06	5.93
<i>All subjects</i>	6.01	6.07	5.96	6.05	6.04	5.94