CHAPTER 40

NEGATION AND THE BRAIN

Experiments in health and in focal brain disease, and their theoretical implications

YOSEF GRODZINSKY, VIRGINIA JAICHENCO, ISABELLE DESCHAMPS, MARÍA ELINA SÁNCHEZ, MARTÍN FUCHS, PETER PIEPERHOFF, YONATAN LOEWENSTEIN, AND KATRIN AMUNTS

40.1. Introduction

This chapter is on the processing cost of sentential negation, and about attempts to uncover its neural basis. We review experimental evidence from normal processing and from aphasia, and discuss its theoretical relevance. Psycholinguistic studies on the processing of negation have been conducted for many years (e.g. Wason 1965; Fodor and Garrett 1966; Just and Carpenter 1971; Kaup and Zwaan 2003; Deschamps et al. 2015). Here, we focus on attempts to localize brain mechanisms entrusted with the processing of negation, and of tests that try to ask whether they are distinct from those mechanisms known to support aspects of syntax. At centerstage, then, will be results of experiments on sentential negation in health and in focal brain disease. As will be seen, there are substantial neurolinguistic hints regarding negation.

In classical logic, negation is a unary connective whose basic function is to reverse the truth value of the (simple or complex) proposition in its scope. In language, things are more complicated (Horn 1989). Negation needn't scope over a proposition, as it can be not only sentential, but also, negate a constituent (cf. *I ate not cereal (but milk)*; *he sat not on the desk, (but under it)*). Indeed, it has been proposed that negation is actually a cross-categorial connective (e.g. Keenan and Faltz 1985). Psycho- and neurolinguistic tests of negation have thus far been largely restricted to sentential negation, and it is therefore the type of negation that we will discuss. We will present experiments and their results from three different sources: first, we'll discuss negation experiments in healthy individuals

whose dependent measure was Reaction Time (RT); next, we'll present similar experiments in which healthy individuals are placed in an MRI machine, and the intensity of the Blood Oxygen Level Dependent (BOLD) response is measured. These two groups of experiments will set the stage for an experiment whose participants were individuals with aphasia subsequent to a focal brain lesion. In this experiment, Error Rate (ER) was the dependent measure.

The chapter is organized as follows: In section 40.2 we review results that indicate that overt sentential negation has a processing cost and brain signature. We note problems in these studies, and move on (section 40.3) to consider implicit, or hidden, negation inside quantifiers that induce downward entailing environments (*more*, *less*, and the like), the use of which we motivate (section 40.4). These considerations lead to section 40.5, where we review recent studies on implicit negation. We discuss behavioral experiments in healthy populations that have shown that this type of negation also incurs a cost, and a study in functional neuroimaging, which localized the processing of negation to a specific brain area in the left anterior insula.

We then move on to negation and negative operators in aphasia (section 40.6). We describe some past experiments, and move on to our experiment with Spanish speaking patients with Broca's aphasia. A discussion regarding the brain localization of negation and negative operators (section 40.7) concludes the chapter.

40.2. Experiments on negation and their design problems

40.2.1. Early experiments

In the mid-1960s, experimental evidence began emerging to the effect that negation makes comprehension more difficult. Fodor and Garrett (1966) were among the first to show that verification times of affirmative sentences are shorter than those of their negated counterparts. Yet, a design problem, inherent in these experiments, seems to have marred the scene: negation is a word, and as such, a direct comparison between a sentence that contains it and one that doesn't is not trivial, for there is always an element in the negated sentence that makes it more complex than its non-negated counterpart. Thus Clark and Chase (1972) took no steps to control for the number of words or syllables, and merely contrasted negative sentences that depict spatial relations between objects with their affirmative counterparts:

(1) a. It is **not** true that the star is above the plus.
    b. It is true that the star is above the plus.

Earlier, Fodor and Garrett explicitly noted that such controls are necessary. They pointed out the difficulty with negation persists even when length and other syntactic factors are controlled (though the type of control remained unspecified). This result has since been

replicated multiple times. Reliable controls, once devised, would warrant the conclusion that indeed, sentential negation incurs cost on the human sentence processing device (see review in Horn 1989: ch. 3 section 2).

However, the nature of reliable controls must be first discussed. We make this point through a short review of recent experiments that sought to uncover brain loci for negation through functional Magnetic Resonance Imaging (fMRI). These have used the affirmative/negative contrast, attempting to solve the problem that negated sentences contain an additional word/morpheme/phonetic object. An affirmative/negative difference in the quantities of the measured variable could be attributed not only to negation, but also to the presence of an extra word, or morphophonological material. Proper controls are not easy to set up, and extant experiments have attempted to do so, but have not been entirely successful.

40.2.2. Recent fMRI studies

We discuss one fMRI study (see Appendix for a review of two recent additional works): Tettamanti et al. (2008) conducted a study which used a clever solution to the control problem described above. It capitalized on the null-subjecthood of Italian: As a sentence may or may not contain a subject, it used a null subject in the negated sentence condition (2a), whereas the overt first person pronoun *io* was used as a control for *non* in the non-negated condition (2b):

(2) a. Adesso non accarezzo il gatto
Now **not** $\text{pet}_{1,\text{Sg}}$ the cat

b. Adesso io accarezzo il gatto
Now **I** $\text{pet}_{1,\text{Sg}}$ the cat

Both sentences had the same number of words and syllables, and were said to differ only in negation. Yet, if (2b) is to serve as a control for (2a), the two sentences must be equivalent not only in number of words and syllables, but also in their meaning, up to negation. Yet this is not the case. In Italian, the use of an overt first-person pronoun serves focus purposes (perhaps more so when subject–verb order is inverted, but focused nonetheless, as Romance subject pronouns need a reason to be overt, especially in the absence of a discourse context as was the case in the experiment). And focus, as is well known, evokes a set of alternatives to the focused element: it is a function that takes a proposition p, a world w, and a set of alternatives A, presupposes that p is true and salient in w, and makes false (or at least less salient) in w every proposition q which is non-weaker than p (Rooth 1992; Fox 2007a). The meaning of (2b), presented informally, is therefore: now I, as opposed to all others among those contained in the context set, am petting the cat. More formally, for the proposition p expressed in (2b), the set of focus alternatives for the first person pronoun I is $A^I = \{x \in D_e/\text{Now x pet the cat}\} = \{$*Now I pet the cat, Now you pet the cat,* . . . $\}$, and focus is the function: $[[\textit{focus}]](A_{<st,t>})(p_{<s,t>}) = \lambda w: p(w) = 1.\forall q \in A: q(w) = 0$. Sentence (2a), which is only uttered with a focused pronoun, therefore means *Now* I *pet the cat, but you don't, and neither does he and neither does she* . . .

Against this background, it is clear that, linguistically, (2b) is not a proper control for (2a), because (2a) contains no focus. The negation difference between the two sentences is confounded with focus. Psycholinguistically, moreover, it is now established that overt first-person pronoun and null pronoun in subject position lead to different processing strategies (Filiaci, Sorace, and Carreiras 2014). Tettamanti et al.'s control is therefore insufficient, and the validity of their conclusions is in doubt, pending further corroboration from better controlled studies. Considerations of a similar nature apply to two other fMRI studies of negation (Bahlmann et al. 2011; Carpenter et al. 1999, see Appendix below for details). An alternative solution is called for.

40.3. A solution: Hidden negation in polar quantifiers

The processing literature harbors a hint for a solution of the control problem of negation experiments. Just and Carpenter (1971) tested the processing cost of negation through a contrast between two polar quantifiers. Following Klima (1964), they invoked a lexical decomposition analysis of polar quantifiers, by which *few = NOT(many)*:

(3) a. Many of the dots are black
b. Few of the dots are red

They devised a sentence verification paradigm, and asked participants to determine the truth value of each sentence against an image that contained only black and red dots (images had 2 red: 14 black or 14 red: 2 black). In such scenarios, (3a–b) have the same truth conditions. They moreover have the same number of words (though not exactly syllables). This analysis, if valid, would indeed get closer to solving the control problem. Under this analysis, the sentences in the pair (3) differ in (almost) a single dimension—one contains a negation whereas the other does not. Indeed, they found a difference: $RT_{few} > RT_{many}$. They claimed that this RT difference is evidence for a process by which *few* decomposes into NOT(*many*).

This perspective, by which certain quantificational expressions contain an implicit negation, whose comprehension incurs a specific processing cost, has been our starting point when we began studying the psycho- and neurolinguistics of implicit negation. But first, we briefly review standard linguistic diagnostics for implicit negation.

40.4. Arguments for implicit negation

Two arguments are standardly invoked in support of the claim that *few* contains a negation. We demonstrate these with the polar pair of proportional quantifiers *more-than-half* and *less-than-half*.

40.4.1. Negation reversal

In classical logic, negation reverses truth value, and the direction of inferences:

(4) a. p is TRUE iff ¬p is FALSE (p = 1 iff = 0)
b. $p \rightarrow q$ iff $\neg q \rightarrow \neg p$

Identical behavior is observed in natural language sentences. Consider (5): the set of boys who are both students and runners is a subset of the set of boys who are students {x/x = boy & x = student & x = runner} ⊆ {x/x = boy & x = student}. While the inference in (5a) is from a subset (student runners) to a superset (students), in (5b) the direction is reversed—from students to student runners:

(5) a. Every boy there was a student and a runner ⇒ Every boy there was a student
b. Every boy there was not a student ⇒ Every boy there was not (both) a student and a runner

The same behavior—reversal of the direction of the inference—persists when *every* and *every . . . not* are replaced with *more* and *less*, respectively (the pair *many*/*few* can be used with the same results). While a visible (or audible) negation is absent, the inference reversal is evidence for its abstract existence:

(6) a. More-than-half of the boys were students and runners
⇒ More-than-half of the boys were students
b. Less-than-half of the boys were students
⇒ Less-than-half of the boys were students and runners

The property of inference reversal is also known as Downward Entailingness, as opposed to Upward Entailingness—the property of maintaining the direction of inferences. A simplified definition of this property is provided in (7):

(7) a. A quantifier Q is Upward Entailing (UE), if A ⊆ A' ⇒ Q(A) ⊆ Q(A')
b. A quantifier Q is Downward Entailing (DE), if A ⊆ A' ⇒ Q(A') ⊆ Q(A)

DE-ness amounts to harboring an implicit negation. Indeed, an abstract negation is considered to be part of *less* (cf. Rullmann 1995b; Heim 2000), which sets it apart from *more*, its control: *more* denotes a relation between two sets of degrees, and the difference between it and *less* is that the latter contains in addition a negation operator. Similar (though not identical) properties are observed for 'negative' verbs, like *surprise*, *doubt*, *deny*, etc. Here, we focus on polar quantifiers.

40.4.2. NPI licensing

Negation licenses Negative Polarity Items (NPIs, Klima 1964; Ladusaw 1979). In (8), the NPI *ever* is licensed by the DE operator *No*, but not by the UE *some*. A similar contrast is observed in (9), except this time, no overt negation is found:

(8) a. No human has ever had a brain transplant
b. #Some human has ever had a brain transplant

(9) a. Less-than-half of the humans have ever had a brain transplant
b. #More-than-half of the humans have ever had a brain transplant

A negation, then, seems to be hidden in *less*, *few*, and other DE quantifiers. Just and Carpenter demonstrated a processing difference between the two, and related it to this negation (for related later experimentation, see Geurts and van der Slik 2005). We call this effect the DE Complexity (DEC) effect. This effect has thus far been demonstrated for quantifiers in subject position, and while it would certainly interact with other quantifier processing effects, for instance those pertaining to Quantifier Raising (e.g. Varvoutis and Hackl 2006), such interactions are outside the scope of the present work, and do not affect its conclusions.

40.5. The processing of hidden negation

40.5.1. Behavioral studies

In a universe of discourse that contains blue and yellow circles, and nothing else, sentences (10a–b) that have an identical number of words and syllables also have the same truth conditions. They do differ, however, in that (10b) contains a negation that (10a) does not:

(10) a. More-than-half of the circles are blue
b. Less-than-half of the circles are yellow

Deschamps et al. (2015) report the results of three speeded verification experiments with polar quantifiers, in which matched auditory sentences were coupled with images that contain blue and yellow circles in varying proportions:

(11) Polar proportional quantifiers:
a. More than half of the circles are blue
b. Less than half of the circles are blue

(12) Polar degree quantifiers
a. Many of the circles are blue
b. Few of the circles are blue

(13) Polar comparative quantifiers
a. There are more blue circles than yellow circles
b. There are fewer blue circles than yellow circles

These experiments were aimed, among other things, at examining the degree to which the DEC Effect generalizes to pairs beyond *many/few*. All three pairs were tested in the same

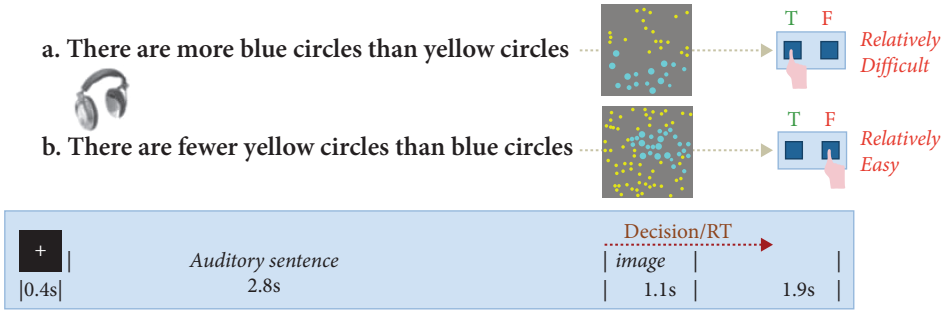

FIGURE 40.1. Form, content, and time-course of stimuli in Deschamps et al. (2015).

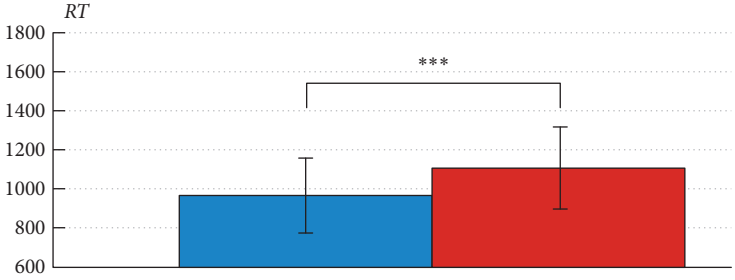

FIGURE 40.2. RTs in msec for UE (**blue**) vs. DE (**red**) stimuli (across proportion and truth value). Error bars mark 1 population SD, Deschamps et al. (2015).

paradigm. Each trial began with a visual fixation point followed by an auditory sentence probe, which was then followed by an image which participants were asked to verify (Figure 40.2). In addition to measuring the processing cost of negation, these experiments tried to see whether DE Cost is affected by properties of the truth-making scenario (in this case, by Weber's Law, Dehaene 1997). Therefore, the blue/yellow proportion in the scenarios was varied along a seven-valued parameter. This proportion determined both truth value (T/F) and task difficulty. That is, each condition contained true and false tokens, and as blue/yellow proportion approached 1, the task became more difficult—an image that contains 16 blue circles and 4 yellow circles is easier to parse than one with 16 blue and 14 yellow ones. This difficulty is famously governed by Weber's Law (cf. Dehaene 1997 and much related literature). In Figure 40.1, for example, we see a more difficult true case and an easier false one.

Three tests were carried out with this verification paradigm. Participants were presented with auditory versions of the sentences in (11)–(13), and were instructed to match the sentence probe with an image as above, and do so as fast as they could.

Two control conditions used visually displayed probes that featured expressions containing inequality symbols (instead of auditory sentences used in the test conditions). We sought expressions whose denotation is similar to the experimental pair, but they are not contrasted by a negation. We used quasi-algebraic expressions that contain inequality symbols (<,>). These symbols are used to compare the size of the values that flank them, and mark the direction of the difference. We used them because their formal definition is characterizable as *less* and *more* respectively, and because each is the converse of each other. Thus instead of proportion-denoting or comparative constructions presented in the auditory modality, probes consisting of colored squares separated by an inequality symbol were presented visually prior to the proportion-depicting image:

(14) (a) ■>■; ■>■
(b) ■<■; ■<■

These conditions were proper controls, as the symbols ">" and "<" have identical geometrical contours, and share no other asymmetry that holds between *less* and *more*, because they merely denote a relation between set cardinalities. Critically, none contains a negation. Therefore, while a negation in *less* reverses the relation denoted by *more*, the two control conditions differ in relation-reversal, yet in the absence of negation.

The experiment thus had a 2 × 2 design, consisting of two contrasts: (i) the Polarity contrast (11)–(13) featured sentences with the quantifiers *more/less*—terms linguistically analyzed as denoting an ordering relation between two sets of degrees, where a reversal is due to a negation operator inside *less*; (ii) the Probe-type contrast (14) featured quasi-algebraic expressions with the symbols >/<—terms denoting an ordering relation, in opposite directions; both symbols are atomic, that is no negation is involved.

Participants were instructed to press the "Match" button in case of a match between the sentence/expression and the image, and "MisMatch" otherwise, and do so as fast as they could. Error rates were low (<15% across all conditions). RTs were recorded, the latter time-locked to image onset as seen in Figure 40.2. In all conditions, the RT functions for the correct responses behaved in keeping with Weber's Law, across all seven values of the proportion parameter and across True and False instances. Figure 40.2 collapses participants' scores across truth value and proportion, and presents a grand mean for each member of the polar pair (11a–b). The DEC Effect is robust (*** = p<.001), manifesting in almost all individual participants.

Returning to our results, it is interesting to note that the RTs for the non-linguistic and the RTs for the quasi-algebraic expressions stood in stark contrast to the pattern obtained for the linguistic materials. No difference between the ">" and "<" conditions was found – $RT_{<} \cong RT_{>}$. This difference between differences manifested as a highly significant Probe type *X* Polarity interaction effect (that is, $[RT_{\textit{less}} - RT_{\textit{more}}] > [RT_{<} - RT_{>}]$).

The two pairs of conditions were presented in different modalities—the *more/less* contrast was auditory, while the >/< contrast was visual. This, however, did not hamper our conclusions, because these were solely based on the *interaction* effect, namely, on the difference-between-differences, which was free of confounds.

In Deschamps et al. and Grodzinsky et al. we drew several conclusions from these results. Of these, two are of interest here: (1) the DEC Effect, found in all instances, was taken to

reflect the processing cost of implicit negation, as this operator was the only one distinguishing between the conditions; (2) the interaction effect was taken to underscore the linguistic specificity of the DEC Effect, as it was restricted to linguistic stimuli.

Further support to these conclusions came from an additional experiment with phrasal comparative constructions (Grodzinsky et al. 2018). Its results were, in fact, used to argue in favor of a decompositional approach to *less*-comparative constructions, whose analysis was a matter of debate in recent years (Rullmann 1995b; Heim 2000; Büring 2007).

40.5.2. A study of hidden negation in fMRI

We were encouraged by the behavioral results, which seemed to see through the processing of DE-ness at an extremely high resolution. This optimism made us turn to the brain, where we hoped to obtain similar findings. We therefore began by pursuing the same issue in a neurolinguistic context—through functional Magnetic Resonance Imaging (fMRI). A preliminary study (Heim et al. 2012) found a neural basis for the DEC effect in the anterior frontal lobe of the left cerebral hemisphere, yet in a better controlled study, we identified a single activation cluster located at the left anterior insula—a deep cortical area adjacent to, but distinct from, Broca's region. This result suggests that the neural basis of the DEC effect is adjacent to Broca's region, but excludes it (Grodzinsky, Agmon, and Deschamps 2016; Grodzinsky, Deschamps, Pieperhoff et al. 2019). The latter study had the same design as Deschamps et al.'s with one difference: it sought a Probe type *X* Polarity interaction effect in signal intensity (rather than RT). Indeed, a single brain area—the left anterior insula—exhibited this interaction effect when the effects of fMRI signal intensity variable were calculated.

These results, which clearly localize core processes of negation, seem to have two important implications. First, they suggest that negation is governed by a brain mechanism that is outside the language areas. Anatomical distinctness suggests functional distinctness: it is quite possible that the operation of negation does not belong to the language module, but rather, to the human logical ability. One can even speculate that if the above is true, then there may be a logic module in the brain. Second, our characterization of the fMRI results helps to derive predictions regarding heretofore unexplored aspects of the deficit in Broca's aphasia: the expectation is that cases in which Broca's region is lesioned, yet the anterior insula is spared, would lead to a subtle comprehension deficit, manifesting as a partial syntactic deficit (e.g. in the style of the Trace-Deletion Hypothesis), with spared negation, and vice versa—we would expect that a lesioned insula and spared Broca's region would result in a pure negation deficit.

These two implications, we note, are contrasted with a commonly held position that language maintains a generic complexity hierarchy, and more complex processes are more costly in time and brain activation, and are the first to break down. What we reported, however, is a fine pattern of selectivity that seems to set syntactic and logical operations apart. In an attempt to broaden our empirical basis, we set ourselves to the hard task of testing these expectations with brain damaged patients. We describe an effort to explore the neural underpinnings of the DEC effect through a study of implicit negation in Broca's aphasia.

40.6. Negation in aphasia

40.6.1. Past experiments with negation in aphasia

Before moving on to our own study, we briefly describe relevant comprehension studies of overt sentential negation with participants suffering from aphasia, who mostly are diagnosed as falling under the Broca's aphasia category. No accurate lesion data are provided, and the diagnosis is mostly based on functional deficits, as detected by clinical tests. These studies have either used a binary-choice selection paradigm (Rispens, Bastiaanse, and Van Zonneveld 2001), in which participants are asked to select the matching picture, or a verification paradigm (Fyndanis et al. 2006), where participants are requested to indicate whether a picture makes a sentence true or false. The results are similar: In Broca's aphasia, performance on negative sentences is overall slightly diminished, but only to a small extent. The overall number of participants in these studies is too small for serious group statistics, but of nine patients tested (coming from several linguistic communities), seven were well above chance on both affirmative and negative sentences, one was at chance on both types, and one performed above chance on affirmatives, and at chance on negatives.

These studies are suggestive: unaffected performance on negated sentences by lesioned patients would corroborate the view that negation—hidden or overt—is not supported by Broca's region. Yet before accepting it, we note that the studies above are lacking in three important respects: (1) They hardly contain any lesion localizing information. (2) They lack clear comprehension scores that can lend credence to the clinical diagnosis. (3) The added lexical complexity of the explicit negation is not controlled. Observing these problems, and in the absence of previous results regarding quantifier polarity in aphasia, we tried to remedy the situation by conducting a study, very much in the spirit of Deschamps et al.'s, with patients for whom we had precise lesion information, as well as comprehension scores that supported the diagnosis.

40.6.2. A new experimental attempt to uncover hidden negation deficits in aphasic patients

We report a preliminary study with patients who were native speakers of Spanish and had suffered stroke, resulting in different diagnoses of aphasic syndromes. Logistical difficulties precluded the recruitment of a large group, and we ended up with six participants (Table 40.1).

Native speaking patient participants were recruited in the Buenos Aires area, through the Instituto de Lingüística, University of Buenos Aires. All participants gave written informed consent in accordance with McGill University's School of Medicine Research Ethics Board, and with the Ethics requirements of the Faculty of Humanities (Filosofía y Letras), the University of Buenos Aires.

Table 40.1. Patient clinical information

Patient	P1 = EC	P2 = RC	P3 = RD	P4 = OV	P5 = MC	P6 = BG
Gender	M	M	M	M	F	F
Edad/ Age	69	59	59	63	58	60
Handedness	Right-handed	Right-handed	Right-handed	Right-handed	Right-handed	Right-handed
Year of Stroke	2001	1997	2000	2003	2012	2009
Schooling	University (17 years)	High school (11 years)	High school (12 years)	University (+17 years)	University (15 years)	University (17 years)
BDAE Picture description (complexity index)	0.6	0.71	0.62	0	1.8	0.77
BDAE Auditory Word comprehension	33,5/37	33,5/37	35,5/37	28,5/37	37/37	34/37
BDAE Order execution	15/15	14/15	15/15	12/15	15/15	15/15
Boston Naming Test	44/60	33/60	46/60	20/60	44/60	34/60
BDAE Sentence comprehension	8/10	10/10	8/10	3/10	10/10	3/10
BDAE word repetition	7/10	10/10	9,5/10	7/10	10/10	10/10
BDAE Repetition of nonsense words	4/5	3/5	4/5	1/5	5/5	4/5
MMSE	28/30	27/30	28/30	28/30	26/30	26/30
Digit span	3	4	5	4	3	3
Reverse digit span	3	3	3	2	2	2
Clinical classification	Broca	Broca	Broca	Transcortical mixed	Anomic	Broca

We tried to overcome the obvious liability inherent in this small sample size by recruiting several current technologies to obtain multiple measures of behavior and brain structure. Below, we describe three behavioral tests and an anatomical study we conducted.

40.6.3. Syntactic comprehension

We began with a commonly used syntactic comprehension battery that tests for a TDH-based deficit in movement-derived constructions (Grodzinsky 1986, 2000), expected in Broca's aphasic patients who suffer from agrammatic comprehension. A sentence-to-picture matching task used "semantically reversible" sentences, featuring a pair of relative clauses with subject- and object-gap, as well as an active/passive pair, each presented concurrent with two images, one depicting a matching scenario, and the other, a mismatch, depicting the same scenario, but with reversed semantic roles (actor ⇒recipient of action and vice versa).

(15) a. **Active** El oso atrapa al mono
The bear catches the monkey

b. **Passive** El mono es atrapado por el oso
The monkey is caught by the bear

c. **Subject Relative** El oso que atrapa al mono es grande
The bear that catches the monkey is big

d. **Object Relative** El mono al que el oso atrapa es grande
The monkey that the bear catches is big

The small number of patients gives little reason to calculate group statistics. Table 40.2 presents raw data (number of correct responses per patient per syntactic type; in parentheses, the number of token sentences per condition):

Of four patients with a diagnosis of Broca's aphasia, three (P1, P3, P6) presented a comprehension performance pattern roughly in keeping with the pattern familiar in Broca's aphasia: above-chance on actives and subject-relatives (where chance-level performance on a binary-choice task is those success rates that are contained within the $p = .95$ confidence interval on a binomial distribution), at-chance levels on passive sentences and object-relative clauses. P4 presented a different pattern—at chance on Subject relatives and below-chance on the Object counterparts). Indeed, he was diagnosed not as suffering from Broca's aphasia, but as a mixed transcortical patient; no comprehension disturbance on this

Table 40.2. Patient performance score on a syntax comprehension task

TYPE	Patient	P1	P2	P3	P4	P5	P6
Active (64)		52	56	61	56	32/32*	48
Passive (64)		36	56	30	38	32/32*	28
Subject Relative (20)		14	19	17	12	20	12
Object Relative (20)		6	11	7	5	20	6

(*For technical reasons, P5 was only tested on half the active and passive sentences)

test was detected for P5, who was diagnosed as suffering from anomic aphasia; the perfect performance of P2 on the passive condition sharply deviated from the expected chance-level pattern. In addition, two patients (P4, P6), though performing lower on Object relative clauses than on Subject relatives, nonetheless deviated from the expected pattern, performing at chance levels on the former and below-chance on the latter (on performance variation in Broca's aphasia see Drai and Grodzinsky 2006a, b).

40.6.4. Polar quantifiers and equivalent symbolic inequalities in verification

We moved to test this small group of patients on implicit negation, using the quantifier polarity test described above, in which Polarity is one factor, and Probe type (linguistic, quasi-algebraic) is the other factor. The final probe design is in Table 40.3 (the Spanish sentences are below the English ones).

All test materials were presented to four neurologically intact control participants, who performed at 92—100% level accuracy on all conditions, indicating the suitability of materials. We then proceeded to the individuals with aphasia. Our hope was to find distinctions that would align with our predictions. The locus found to be activated by DE-ness in the fMRI study of healthy participants was the left anterior insula. We therefore expected that patients whose lesions included this region would produce high error rates on conditions that contain DE quantifiers (*few*, *less-than-half*), and hopefully low error rates on their UE counterparts (containing *many*, *more-than-half*), as well as on the conditions with non-linguistic probes.

As our participants were stroke victims, we made two minor modifications in Deschamps et al.'s experiment, to make the task easier for the patients: (1) Trial overall duration was extended from 6.4 sec to 12 sec, and the image stayed on the screen from its onset to trial's end. All else was the same as in Deschamps et al. (2) The number of test items was reduced: from 7 proportions per condition type, we moved to 5, therefore, the total number of trials was $5_{\text{proportions}}*8_{\text{tokens}}*2_{\text{truth-values}}*6_{\text{conditions}} = 480$ trials. (3) The verification task was not speeded, and error-rate (not RT) was the dependent variable. These modifications were global, ranging across all 6 conditions. Participants were tested in

Table 40.3. The six conditions of the experiment, organized by factor

		Probe type		
		Linguistic		**Non-linguistic**
Polarity	***UE***	More-than-half of the circles are blue/yellow *Más de la mitad de los círculos son celestes/amarillos*	Many of the circles are blue/yellow *Muchas de los círculos son celestes /amarillos*	🟨 > 🟦 ; 🟦 > 🟨
	DE	Less-than-half of the circles are blue/yellow *Menos de la mitad de los círculos son celestes/amarillos*	few of the circles are blue/yellow *Pocos de los círculos son celestes /amarillos*	🟦 < 🟨 ; 🟨 < 🟦

their homes in multiple 30 minute sessions. Our dependent variable in this study was error rate (proportion correct).

The results (Figures 40.3–40.4), presented as individual means per condition (collapsed across proportion and truth value), reveal a mixed picture.

These results can be succinctly described as follows:

a. Two patients failed to exhibit a selective pattern—P5 was near ceiling on all conditions, whereas P6 was at chance across-the-board (where chance-level performance on a binary-choice task is those success rates that are contained within the p = .95 confidence interval on a binomial distribution, or 7–11 successful trials out of 16). P6 was also at chance level on both subject and object relative clauses. Curiously, their production patterns, as it emerges from their clinical scores on language production tests, are nearly identical to one another.
b. For P1–P4, performance on UE quantifiers and on both symbolic inequalities was above-chance level.
c. For P1–P4, the remaining four patients, overall performance on UE quantifiers generated fewer errors than performance on their DE counterparts, with one odd exception (P3: chance level on *more-than-half*, above-chance on *less-than-half*).
d. Fourth, performance on DE quantifiers varied greatly. This variability made group statistics superfluous. Nevertheless, these behavioral data seem to be telling us a fairly clear lesson: of the three types of relational expressions used in this experiment, those containing DE quantifiers are the most vulnerable. Next, we tried to relate the behavioral deficit to lesion anatomy, through a detailed study of the patients' lesions.
e. P2 stands out when individual performance patterns are examined. His syntax comprehension scores were good except the Object relative clauses. On the present test, his performance was near-normal on the UE quantifiers and on both symbolic inequalities, yet he performed *below* chance (1–2 successes of 16 trials) on linguistic conditions—those containing an implicit negation in a DE quantifier. His performance thus indicates that he interprets *less-than-half* as *more-than-half* and *few* as *many*. While only observed for one patient at present, this pattern is reminiscent of results from language acquisition (Clark 1970, *passim*).

The main result, which we discuss below, is the tendency to fail on linguistic DE conditions, and the lack thereof in the symbolic conditions. Regarding performance patterns, these are intricately variable and elude an immediate explanation. It seems that greater numbers of patients are required for any firm conclusion to emerge. We now turn to the anatomical side of this study.

40.6.5. Lesion anatomy

All our participants received a brain scan. Subsequently, their lesions were masked manually, in order to allow for precise anatomical localization and analysis, through the use of the probabilistic, cytoarchitectonic JuBrain atlas, which contains maps of cortical areas and subcortical nuclei as defined in a sample of ten postmortem brains (Amunts and Zilles 2015). The JuBrain atlas carries cortical maps based on cytoarchitectonic analysis and computational methods

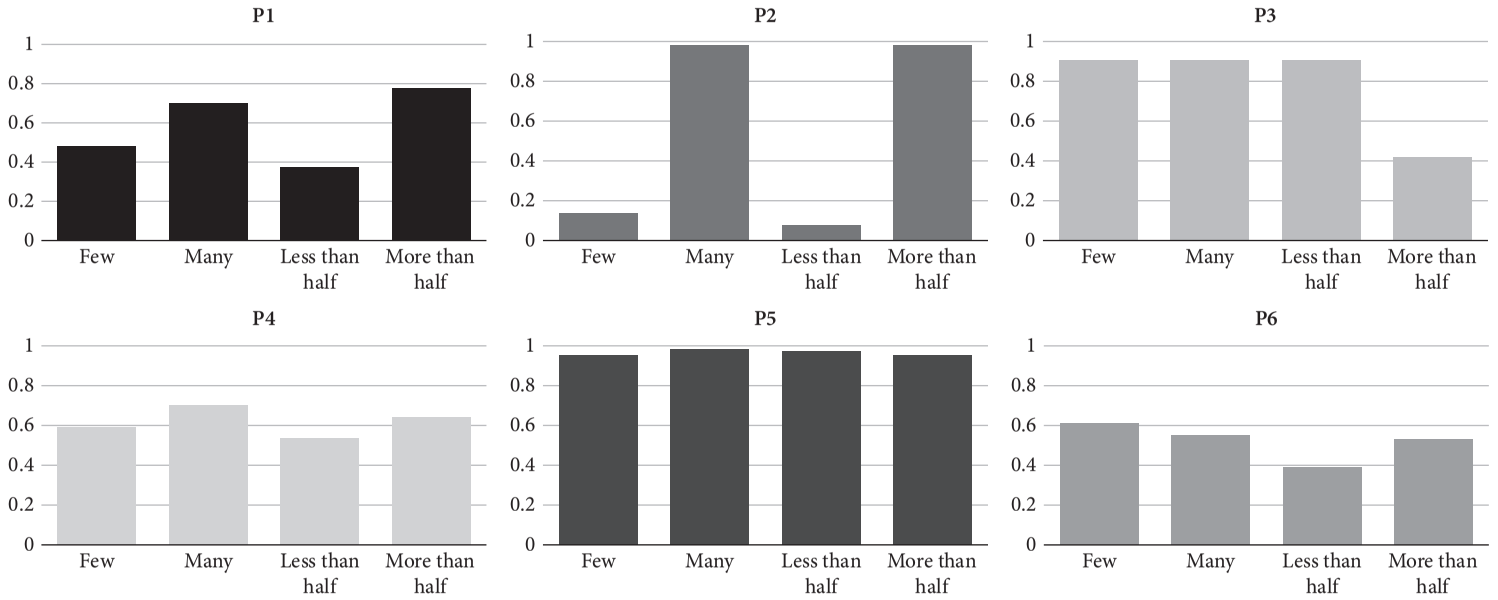

FIGURE 40.3. Percent correct on each condition per patient—the linguistic conditions.

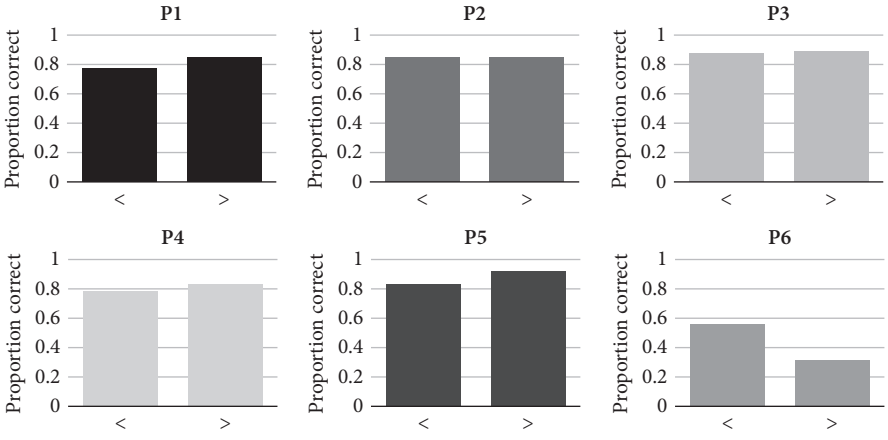

FIGURE 40.4. Percent correct on each condition per patient—the non-linguistic conditions.

including image analysis and statistical analysis that parse the brain's grey matter into areas. While some of the maps, in particular of primary sensory and motor areas, seem to be similar to what are historically known as Brodmann Areas (e.g. BA4, BA17, and, to a large extent, BA44 and 45 of Broca's region; Amunts et al. 1999), the vast majority of areas do not have similar counterparts in this historical map (Brodmann 1909), but provide a more detailed subdivision of the cerebral cortex. In contrast to Brodmann's map, the JuBrain atlas considers intersubject variability as a feature to describe an area ("probabilistic"), and provides true stereotaxic information, a prerequisite for comparison with findings from functional neuroimaging. The JuBrain atlas has a spatial resolution of $1mm^3$ voxels. It is freely available to the research community and linked to other data modalities in the HBP Human Brain Atlas (<https://www.humanbrainproject.eu/en/explore-the-brain/>).

This atlas allows computational comparisons between distinct brain areas, and is a tool to evaluate voxels, or clusters of voxels, acquired by other methods in healthy subjects and patients, for example fMRI activation clusters, voxels containing focal lesions, etc. Such methods produce results whose quality is quite different from visual inspection of topographic landmarks observed in an image. When the coordinates in common reference space of a cluster's voxels are known, they can be located by co-registration to the *JuBrain* atlas (cf. Amunts et al. 2004; Santi and Grodzinsky 2012 for applications to fMRI language studies). Once a lesion is mapped (or masked through a difficult semi-automatic process), the anatomical addresses of all its voxels are known, and can be mapped onto the atlas (Hömke et al. 2009), resulting in information about the cytoarchitectonic correlates of the lesion. As lesions are caused by pathological processes that do not respect histological boundaries, we typically obtain a list of cytoarchitectonic areas that are lesioned, where each is listed with the degree to which it is compromised. Thus in Table 40.4, 76% of the posterior part of Broca's area of the left cerebral hemisphere, namely area 44L in the JuBrain, is lesioned for patient P1; while 53% of the posterior part of his left Broca's area, 45L, and only 9% of his left anterior insula Id7_L, are lesioned. These sophisticated mapping tools therefore provide a quantitative picture of the patient's brain, which can be compared to his/her impaired functions for localizing purposes.

Table 40.4. Percent lesion in four cytoarchitectonic language areas

		P1	P2	P3	P4	P5	P6
44L	(Broca's region)	76	94	85	94	10	94
45L	(Broca's region)	53	94	66	85	5	80
Id7_L	(anterior Insula)		99	100	100	4	52
TE3_L	(Wernicke's area)	4	13	5	97	1	95

Clearly this atlas becomes amenable to computational investigations when data from larger numbers of patients are available. The extreme difficulty in patient recruitment and scanning has left us with only six patients, all affected in the left hemisphere. While every one of these patients is also affected in other brain regions, we restricted ourselves to four language regions. In Table 40.4, we show the extent of the lesion (% destroyed) each of them suffered in four left hemispheric brain areas, known to support language. In the absence of a larger sample, we can only use this information informally to find some generalities as well as individual differences:

40.6.6. An informal analysis

We have thus far reported four different classes of quantitative measures about our patients: clinical, anatomical, syntactic, and semantic, negation-related. Ideally, we would apply computational methods in order to uncover systematic relations between these measures. However, the paucity of cases provides little opportunity for quantitative analyses, limiting us to an informal discussion.

Consider first the uninformative performance patterns of P5 and P6: Table 40.4 seems to explain them. The relevant anatomical areas in the brain of P5 are virtually unaffected, hence his performance is indeed expected to be near-normal; the brain of P6, by contrast, is affected in a sweeping fashion—both anterior and posterior language regions, including Wernicke's area, are compromised. As Table 40.2 shows, one cannot be sure that P6 even understood the instructions. Her across-the-board chance-level performance is thus not unexpected.

Next, consider the performance patterns of P1 and P2. P1's syntactic comprehension is typical of Broca's aphasia, whereas P2's is better (though not far from typical). In the polarity experiment, they were both well above chance on the symbolic conditions and on the UE linguistic conditions. On the DE conditions, P1 was at chance whereas P2 was *below* chance. For both, Broca's region is seriously injured. Yet, whereas P2 lost his left anterior insula, 91% of this brain area is spared for P1. It is possible (though by no means certain) that this anatomical difference accounts for their performance difference. Still, any assertion that the difference in these patients' lesions translates directly into the measured performance difference would require a broader empirical basis, namely many more patient scores.

Less clear is the relation between the lesion and the behavioral patterns of the remaining two patients: for P3, who presents a typical picture of Broca's aphasia, the anterior language region is

destroyed, while the posterior one—Wernicke's area—is spared. By contrast, for P4, all four language areas are almost completely wiped out. Still, he seems to present with a structured pattern, performing above-chance on both symbolic conditions, (barely) above-chance on the conditions that contain a UE quantifier, and at chance level on the DE conditions.

40.7. Implications

We reviewed experimental results regarding negation-related behavior in the healthy and impaired brain. We began by arguing that tightly controlled experiments with overt negation are hard to come by, and proposed an alternative method: implicit negation, hidden inside proportional, degree, and comparative quantifiers. We provided linguistic methods for the detection of this negation and showed that it incurs a processing cost—the Downward Entailing Cost or DEC effect, found in several RT experiments with a variety of implicit negations. This finding supports a decompositional approach to implicit negation generally, but specifically, in the realm of comparative constructions.

The angle we proposed was wider, because our interests lie not only in the representation and processing of negation, but also its brain mechanisms. We briefly reviewed findings from imaging suggesting that (i) negation is localized in cortex; (ii) the neural tissue that support mechanisms for the processing of negation are distinct from, though adjacent to, Broca's region. We then proceeded to describe the details of an experiment with Spanish speaking aphasic patients, that sought to establish similar conclusions through the use of five data sources: clinical diagnosis, syntax comprehension test, a hidden negation test, and anatomical lesion analysis. Here things became more complicated. The behavioral pattern we uncovered was not inconsistent with the expected one, but vague at times. Yet, while the performance patterns observed through RT and fMRI studies in health were very refined, we noted high inter-patient variability in our population, one that resulted in a coarse anatomico-behavioral pattern containing apparent contradictions. The range of data for each patient were broad, yet the number of cases was too small for firm conclusions. The limited evidence we obtained from aphasia, then, is suggestive, though not compelling. At present, the failure to find a stable relation between behavioral deficit and lesion anatomy is due to the small number of patients tested. In the past, it has been shown that patterns emerge with larger number of patients (Drai and Grodzinsky, 2006a, b). We therefore hope that methods for large-scale testing of brain-damaged patients will be developed, to enable a more solid lesion-based perspective on negation processing and related cognitive components.

Acknowledgements

Partially supported by Israel Science Foundation, Grants No. 2093/16 (Y.G.) and No. 757/16 (Y.L.), The Gatsby Charitable Foundation (Y.L.) and by the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 7202070 (HBP SGA1) and No. 785907 (HBP SGA2). Contact: yosef.Grodzinsky@mail.huji.ac.il.

APPENDIX

More on the control problem of negation

An early study (Carpenter et al. 1999), carried out with materials developed in psycholinguistics (Clark and Chase 1972) took no steps to control for the number of words, and merely contrasted activation for negative sentences that depict spatial relations between objects, and their affirmative counterparts:

(A) a. It is **not** true that the star is above the plus.
b. It is true that the star is above the plus.

Later fMRI studies probed negation while attempting to set up proper controls. In a 2 × 2 study of double negation with complex sentences, Bahlmann et al. (2011) similarly controlled the appearance of a negation word with another. The German *nicht* was featured in the matrix clause and/or in an embedded clause. When in the former, it was controlled by *schon* (which the authors translated as *indeed*), whereas in the embedded clause it was controlled by *wirklich* (translated as *really*):

(B) a. Es ist **nicht** wahr, dass Peter Thomas letzte Woche für das Projekt **nicht** einstellte.
It is not true, that Peter did not hire Thomas last week for the project.
b. Es ist **schon** wahr, dass Peter Thomas letzte Woche für das Projekt **wirklich** einstellte.
It is indeed true, that Peter really hired Thomas last week for the project.

Like Tettamanti et al.'s study, Bahlmann et al. feature expressions that carry non-negligible semantic weight (*schon* and *wirklich* introduce a host of presuppositions). Here, too, semantic equivalence is not maintained.