Journal of Semantics, 41, 2024, 77–101
https://doi.org/10.1093/jos/ffae003
Advance Access Publication Date: 23 April 2024
Article

The Domains of Monotonicity Processing

I-An Tan
Neurolinguistics Lab, Edmond & Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, 9190401, Israel

Nir Segal
Department of Linguistics, LLCC, The Hebrew University of Jerusalem, Jerusalem, 9190500, Israel

Yosef Grodzinsky
Neurolinguistics Lab, Edmond & Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, 9190401, Israel

Institute of Neuroscience and Medicine (INM-1), Research Center Jülich, Jülich, 52428, Germany

First version received 14 May 2023; Second version received 13 February 2024; Accepted 15 February 2024

Abstract

This paper reports an investigation into the nature of Negative Polarity Item (NPI) licensing conditions from a processing perspective. We found that the processing cost of Downward Entailingness (*a k a* the *Monotonicity Effect*) is determined by the number of monotonicity reversals of NPI domains, rather than by the number of Downward-Entailing (DE) operators. This conclusion is not based on the standard judgment paradigm, but rather, on the measurement of continuous variables (error rates, Reaction Times (RTs)) in a verification task, in which the truth value of a sentence is determined against a scenario. We conducted two experiments with sentences that contained one or two DE operators, which featured in different syntactic configurations. We explored how RT is affected by the manipulation of both the number of DE operators, and the syntactic environments in which they reside. We ran these experiments in Hebrew and in English, with different participant populations and different testing methods. Despite the linguistic subtlety of the theoretical issues, our results were remarkably sharp, leading to two firm conclusions: (i) that processing time is determined not by the number of DE operators, but rather, by the monotonicity of the minimal constituent in which they reside; (ii) that DE-ness is not a property of operators, but of environments. We show how our results bear directly on the current debate about the nature of monotonicity, which we describe below. Finally, we provide quantitative tests of alternative, non-semantic explanations, and show how our results do not support them.

© The Author(s) 2024. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

1. INTRODUCTION

This paper presents the results of two experiments that provide new insight into the debate on the licensing conditions of Negative Polarity Items (NPIs). Specifically, we approached this issue from the perspective of processing, and asked whether monotonicity has measurable processing consequences. To this end, we manipulated the number of Downward-Entailing (DE) operators as well as the syntactic environments in which they reside, and measured Reaction Time (RT) in a verification task.

NPIs such as *any* and *ever* in English have interested linguists since the 1960s (Klima, 1964; Baker, 1970; Fauconnier, 1975; Ladusaw, 1980; Gajewski, 2011; Xiang et al., 2009, 2013, 2016; *passim*). These items must be licensed by some sort of negativity—a negative marker, such as *no* or *not*, or a DE operator, such as *less* and *few*.[1] Hence, the unlicensed NPIs in the sentences in (1a,c,e) result in ungrammaticality, but sentences (1b,d,f) are grammatical, as their NPIs are licensed by a negation (1b) or a DE quantifier (1d,f).

(1) a. *The kids ate *any* cookies at the birthday party.
b. The kids did not eat *any* cookies at the birthday party.
c. *More than 3 kids ate *any* cookies at the birthday party.
d. Less than 3 kids ate *any* cookies at the birthday party.
e. *Many kids ate *any* cookies at the birthday party.
f. Few kids ate *any* cookies at the birthday party.

Whereas it is easy to see that an NPI requires a licensor, to formulate the precise licensing conditions is more difficult. Everyone agrees that a DE operator is needed to license an NPI, yet the question is how the licensor and licensee are related. Homer (2021) divides theories of NPI licensing into two groups: an *Operator-Based Approach* (OpBA), by which all an NPI needs is to be in the scope of a DE operator, as stated in (2), and a stricter *Environment-Based Approach* (EnvBA), by which the environment needs to be DE, as in (3).

(2) Operator-Based Approach (OpBA): An NPI is licensed only if it is in the scope of a DE expression. (Fauconnier, 1975; Ladusaw, 1980).

1 Entailment, downward-entailing function and downward-entailing environment are defined as follows (for concreteness, we adopt Crnič's (2014) definition here):
(i) Cross-Categorical Entailment ($\Rightarrow$)
a. For p, q which are truth values: $p \Rightarrow q$ iff $p = 0$ or $q = 1$;
b. For f, g of type $\langle\sigma,\tau\rangle$: $f \Rightarrow g$ iff for all x of type σ, $f(x) \Rightarrow g(x)$.
(ii) A Downward-entailing function:
A function f is DE iff for any x and y in the domain of f such that $x \Rightarrow y$, $f(y) \Rightarrow f(x)$.
(iii) Downward-entailing environment:
A constituent X is DE with respect to a sub-constituent Y of type α iff replacing Y with a variable of type α and binding it by a λ-abstractor adjoined as a sister of X yields a DE function.
An Upward-entailing (UE) function and UE environment are defined symmetrically:
(iv) UE function:
A function f is UE iff for any x and y in the domain of f such that $x \Rightarrow y$, $f(x) \Rightarrow f(y)$.
(v) UE environment:
A constituent X is UE with respect to a sub-constituent Y of type α iff replacing Y with a variable of type α and binding it by a λ-abstractor adjoined as a sister of X yields a UE function.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

(3) Environment-Based Approach (EnvBA): An NPI α is licensed in sentence *S* only if there is a constituent A of *S* containing α such that A is DE w.r.t the position of α. (Gajewski, 2005).

At issue is the domain of an NPI (Monotonicity Domain henceforth). The OpBA claims that it is the scope of a DE operator. Namely, whether or not it is licensed depends solely on its structural relation to that operator. Nothing else is relevant. This approach is contrasted with the EnvBA, which argues that the Monotonicity Domain that licenses an NPI is determined by the DE-ness of the *syntactic environment* in which the NPI resides. This type of DE-ness, moreover, may not be necessarily computed in whole sentences, but rather, in constituents that make up the syntactic environment necessary for licensing. As our experimental inquiry seeks to adjudicate between the OpBA and the EnvBA empirically, we focus on specific syntactic considerations, carefully selected out of the dense literature on the licensing conditions. We begin with empirical arguments that rely on plausibility and acceptability judgments in the context of two DE operators. Here, a sentence is needed, in which an NPI is in the scope of a DE operator, but at the same time, the relevant environment is non-DE. If the sentence is acceptable, the DE-ness of that environment may not matter, as the OpBA would have it; yet if the sentence is unacceptable, then the DE-ness of the environment does matter, as the EnvBA suggests.

Homer (2021) provides evidence from French in favor of the EnvBA. He first presents evidence that *quoi que ce soit* (= anything) is a (weak) NPI, licensed by a negative adjective (4) or a negation (5):

(4) Il est **im**possible que Jean ait fait *quoi que ce soit* pour aider la Mafia.
it is impossible that Jean have.SUBJ done what that this be.SUBJ to help the Mafia.
'It is impossible that Jean did anything to help the Mafia.'

(5) Il **n'** est **pas** possible que Jean ait fait *quoi que ce soit* pour aider la
It NE is NEG possible that Jean have.SUBJ done what that this be.SUBJ to help the Mafia.
Mafia.
'It is not possible that Jean did anything to help the Mafia.'

The acceptability of (4)–(5) is expected by the OpBA, because both contain a DE operator that can serve as licensor. If so, then the presence of *both* licensors must certainly license the NPI. Yet Homer points out that in "flip-flop" sentences, an NPI is anti-licensed:

(6) *Il **n'** est **pas** **im**possible que Jean ait fait *quoi que ce soit* pour aider
It NE is NEG impossible that Jean have.SUBJ done what that this be.SUBJ to help
la Mafia.
the Mafia.
'It is not impossible that Jean did anything to help the Mafia.' (Homer, 2021)

Homer shows how this unacceptability follows from the EnvBA. The intuition is that one cannot find in this sentence a DE environment that contains *quoi que ce soit* and only one of ***pas*** and ***im**possible* (which are both in the same TP). These two operators (each individually licensing the NPI) are coupled in this configuration, and thus NPI licensing fails. Homer proposes that this failure is because licensing here may only be done by the node that dominates both of them, which is Upward-Entailing (UE), as the combination of the two DE operators ***n'...pas*** and ***im**possible* leads to UE-ness. Therefore, *quoi que ce soit* resides in a UE environment, and unacceptability follows.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

The EnvBA succeeds because it requires that licensing is dependent on a DE environment, not a DE operator. Homer calls this environment "the domain of an NPI":

(7) Domain of an NPI: A constituent γ which contains the NPI π is a domain of π if and only if the acceptability of π can be evaluated in γ.[2] (Homer, 2021)

(8)–(11) are the configurations at issue, their actual acceptability, and the predictions of each approach in keeping with its definition of Monotonicity Domain. Below is a presentation in which γ marks the Monotonicity Domain of an NPI according to the EnvBA (⇧ denotes a UE constituent and ↓ denotes a DE one):

			Acceptability Prediction	
			OpBA	EnvBA
(8)	* … [$_{\gamma\Uparrow}$ <u>UE</u>… ***NPI*** ….]	0*DE licensor in γ	*	*
(9)	✓ … [$_{\gamma\downarrow}$ <u>DE</u>… ***NPI*** ….]	1*DE licensor in γ	✓	✓
(10)	* … [$_{\gamma\Uparrow}$ <u>DE DE</u>… ***NPI*** ….]	2*DE licensors in γ	✓	*
(11)	✓ … [$_{\Uparrow}$ <u>DE</u>… [$_{\gamma\downarrow}$ <u>DE</u>…***NPI*** ….]	1*DE licensor in γ	✓	✓

But what is γ? Homer proposes that a Polarity Phrase (PolP) is the smallest, but not the only, domain of *quoi que ce soit*. Since (6) is ungrammatical, *n'…pas* and *impossible* must be in the same domain there; on the other hand, in the acceptable (12a), the configuration contains a domain with the NPI and a single licensor. He further points out that in (12b), both the conditional-*if* and *impossible* are DE, and *quoi que ce soit* is nonetheless licensed.

(12) a. Il est **im**possible que Jean **n'** ait **pas** fait *quoi que ce soit* pour
It is impossible that Jean NE have.SUBJ NEG done what that this be.SUBJ to
aider la Mafia.
help the Mafia.
'It is impossible that Jean didn't do anything to help the Mafia.'

b. **S**'il est **im**possible que Jean ait fait *quoi que ce soit* pour aider la mafia, je lui présenterai mes excuses.
'If it is impossible that Jean did anything to help the Mafia, I will apologize to him.'

He concludes that these two DE operators cannot be in the same minimal domain, and therefore, there must be a domain below CP, but above VP, where the NPI *quoi que ce soit* is licensed. This domain is PolP.[3] Homer's LFs of (4)–(6) and (12a-b) are shown below. They

2 Homer (2021) defines *acceptability* of an NPI as: an NPI π is acceptable in a constituent γ if and only if γ has the appropriate monotonicity w.r.t. the position of π in γ.

3 Homer (2021) proposes that TP is another domain of NPI in French, given that the following sentence, which contains two DE operators—conditional-*if* and 'at most five', is grammatical. Assuming that 'at most five' sits in [Spec, TP] and conditional-*if* sits in [Spec, CP], there must be a domain containing only 'at most five' but not conditional-*if*, namely, TP.
(i) Si au plus cinq personnes ont fait quoi que ce soit pour aider la Mafia, nous sommes sauvés.
'If at most five people did anything to help the Mafia, we are good.'

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

feature the relevant domains (i.e., PolP and TP), and annotated for their monotonicity with respect to *quoi que ce soit*:

(4)' ...[$_{TP\downarrow}$ T [$_{PolP\downarrow}$ **im**possible [$_{CP}$ que Jean T [$_{PolP\Uparrow}$ [*quoi que ce soit*]$_1$ faire t$_1$]]]].
(5)' ...[$_{TP\downarrow}$ T [$_{PolP\downarrow}$ **pas** possible [$_{CP}$ que Jean T [$_{PolP\Uparrow}$ [*quoi que ce soit*]$_1$ faire t$_1$]]]].
(6)' *...[$_{CP}$ [$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ **pas im**possible [$_{CP}$ que Jean T [$_{PolP\Uparrow}$ [*quoi que ce soit*]$_1$ faire t$_1$]]]].
(12a)' ...[$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ **im**possible [$_{CP}$ que Jean T [$_{PolP\downarrow}$ **pas** [*quoi que ce soit*]$_1$ faire t$_1$]]]].
(12b)' ...[$_{TP\Uparrow}$ [$_{CP\Uparrow}$ **Si** T [$_{PolP\downarrow}$ **im**possible [$_{CP}$ que Jean T [$_{PolP\Uparrow}$ [*quoi que ce soit*]$_1$ faire t$_1$]]]] T].

This perspective on a Monotonicity Domain enables Homer to derive a more refined version of the Environment-based licensing condition for NPI:

(13) Environment-based Licensing Condition for NPIs (refined): An NPI $\boldsymbol{\alpha}$ is licensed in sentence *S* only if $\boldsymbol{\alpha}$ has a DE domain in *S*.

Homer's data support the EnvBA. An operator-based approach cannot explain why (6) is unacceptable, as the NPI is in the scope of a DE operator. Note that Homer's discussion, while pertinent to our specific experimental inquiry, addresses few among the many issues in the analysis of NPIs. A proper analysis must involve semantic and pragmatic assumptions, which, together with syntactic ones, conspire to account for the richly complex array of monotonicity-related phenomena as revealed through the distribution of NPIs (cf. Bar-Lev & Fox, 2020; Chierchia, 2013; Crnič, 2019; Guerzoni & Sharvit, 2007; von Fintel, 1999; Gajewski & Hsieh, 2014, to mention just a few). The phenomena on which Homer and the present paper focus mostly call for a syntactically-constrained analysis of monotonicity, by pitting the OpBA and the EnvBA against one another.

2. THE PROCESSING COST OF DOWNWARD-ENTAILINGNESS

Here, we investigate the nature of DE-ness from a processing perspective. The empirical basis for hypotheses regarding NPI licensing is typically grammaticality/acceptability judgments, whereas we seek to predict, and then precisely measure, the processing cost of monotonicity. Below, we construct a mapping between representation and processing, from which we can then derive explicit predictions about measurable processing costs that the DE-ness of a string incurs, and the manner in which they are dictated by each of the grammatical approaches to monotonicity. In doing so, we will expand the range of empirical evidence that bears on the definition of NPI licensing conditions. As extant evidence is presently limited to patterns of NPI licensing, we now show that results from experiments with continuous variables, mainly Reaction Times (RTs), provide a sharp picture that bears on this issue directly.

Our starting point is the well-known *Monotonicity Effect* in processing: main clauses that contain a DE operator (but no filler-gap dependencies) take longer to process than their UE counterparts (e.g., Just & Carpenter, 1971; Clark & Chase, 1972; Tian et al., 2010; Deschamps et al., 2015; Tian & Breheny, 2016; Agmon et al., 2019, 2022; Dudschig et al., 2019; Schlotterbeck et al., 2020; Wang et al., 2021). Many, if not most, of these experiments control for psychological, morphophonological, and syntactic confounds, which strongly suggests that the monotonicity effect is *not* due to frequency, length or generic complexity (cf. Deschamps et al., 2015; Grodzinsky et al., 2021; Agmon et al., 2022 for recent discussions). Yet robust as this result may be, it is not easy to interpret. A thought that most naturally

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

comes to mind is that *the parser never expects DE-ness*. Its encounter with DE-ness therefore forces a costly monotonicity reversal:

(14) *Monotonicity Processing Hypothesis* (MPH):
a. The parser never expects DE-ness (as defined by the theory of NPI licensing)
b. Each monotonicity reversal incurs a processing cost.

The MPH derives the *Monotonicity Effect*. The idea, to be clear, is that the theory of NPI licensing dictates the processing of monotonicity, for which the parser has a default UE expectation (14a); a monotonicity reversal incurs a measurable cost (14b). The question here is how the monotonicity reversal is counted? MPH is expected to pit the OpBA and EnvBA against one another in terms of their different definitions of the notion "Monotonicity Domain". Thus (14), coupled with each of these accounts, has clear cost predictions. Below, we report experiments that sought direct experimental support for it. We hoped, that is, to obtain evidence to adjudicate between the two theories—the OpBA and EnvBA. Coupling the MPH with the former (i.e., MPH + OpBA) predicts that the cost is determined by the number of DE operators in the incoming sentence; by contrast, MPH + EnvBA predicts that the processor is taxed by the number of Monotonicity Reversals of Domains (MRDs) of an NPI.

We therefore investigated the conditions under which DE-ness incurs a processing cost. We measured continuous variables—mainly Reaction Time (RT)—and explored their relation to the number of DE operators in the incoming sentence, as well as to the domain in which these operators are contained. When the OpBA is mapped onto RT (as an index of processing cost), it predicts RT to grow with the number of DE operators, as each monotonicity reversal is said to be costly; the EnvBA also predicts RT growth with the number of monotonicity reversals, but only if these reversals occur in distinct NPI domains. The table below details these predictions as the sign of ΔRT between members of minimal pairs of sentences, in which DE-ness and domain size are varied systematically (γ and ß denote the domains of an NPI):

		Processing Cost ($\Delta RT = RT_b - RT_a$)	
		OpBA	EnvBA
(15) a. … [$_{\gamma}$⇧ UE….]	1*UE in γ	+	+
b. … [$_{\gamma}$↓ DE ….]	1*DE in γ		
(16) a. … [$_{\gamma}$↓ DE UE….]	1*UE, 1*DE in γ	+	-
b. … [$_{\gamma}$⇧ DE DE ….]	2*DE in γ		
(17) a. … [$_{ß}$↓ DE… [$_{\gamma}$⇧ UE ….]	1*UE in γ	+	+
b. … [$_{ß}$⇧ DE… [$_{\gamma}$↓ DE ….]	1*DE in γ, 1*DE in ß		

As noted above, the contrast in (15) has been tested extensively in main clauses (e.g., *few/many of the dots are red*, Just & Carpenter, 1971), resulting in a positive ΔRT, denoted by "+", illustrating to the *Monotonicity Effect*. These sentences feature a single DE/UE operator contrast, and likewise one DE/UE contrast within the domain γ. Thus, both approaches to NPI licensing (coupled with the MPH) predict it.

Equipped with the MPH and a positive ΔRT_{15}, we sought to distinguish between the two approaches to NPI licensing by measuring the effects of processing two operators in a single domain γ, as depicted in (16). Homer's flip-flop effect (6) suggested that while (16a)

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

is DE, (16b) is UE. As the table shows, the EnvBA+MPH predicts an RT flip-flop as well—reversal in the sign of ΔRT_{16} relative to ΔRT_{15}. Thus, negative ΔRT_{16} is expected (denoted "-" above), as RT_{16a} should be greater than RT_{16b}. A ΔRT sign reversal between (15) and (16) has actually been found by Grodzinsky et al. (2021) and Tan et al. (2023). The EnvBA would appear to be vindicated, as the OpBa predicts the opposite, namely that the processing of two DE operators would be more taxing than one, regardless of where they stand relative to $\boldsymbol{\gamma}$. Yet the sign reversal found for (the negative) ΔRT_{16} relative to (the positive) ΔRT_{15} stops short of distinguishing between the two approaches to NPI licensing: the sentences in (16) contain more operators (and more words) than (15). A reversal in the sign of ΔRT may be due to that. Further controls are missing. The remedy comes in the form of (17), that has the very same words as (16), yet the two operators are in different domains—one in $\boldsymbol{\gamma}$, and the other, in ß. If the (15)–(16) sign reversal is due to added operators, then it should persist in (17); this is what OpBA would predict. Yet if the negative ΔRT_{16} is due to flip-flop in $\boldsymbol{\gamma}$, it should reverse to positive in (17), because the two DE operators are in separate domains. Overall, then, OpBA and EnvBA diverge in prediction, as the table indicates. Thus (15)–(17) provide the full array of sentence types, and feature as stimuli in our experiment.

3. EXPERIMENTAL METHODS

3.1. *A speeded sentence-picture verification task that measures processing cost*

In the current study, we attempt to adjudicate between the OpBA and the EnvBA by providing quantitative processing evidence from two experiments: one is a lab experiment with native Hebrew speakers, and the other is an online experiment with native English speakers. We used the common speeded sentence-picture verification task (SSPVT) paradigm. In this task, participants first hear a sentence, and then see an image that appears immediately thereafter.[4] The SSPVT requires them to indicate, as fast as they can, whether the image makes the sentence true or false.

We used two DE operators, *not* and *less*, to construct two different structures of double negation: the *intra-domain structure* and the *cross-domain structure*. In the *intra-domain structure*, there is no domain that contains only one DE operator. Namely, both of the eligible domains contain either zero or two DE operators (e.g., *Not less than half of the circles are blue*). In the *cross-domain structure*, there is one domain which contains only one DE operator (e.g., *Less than half of the circles are not blue*). We measure RTs for these, as well as for their *more* baseline counterparts, in the SSPVT. To forecast, in both experiments, participants responded faster to the conditions with two intra-domain DE operators than the conditions with two DE operators across-domains. This is predicted by MPH + EnvBA but not the OpBA.[5]

4 For the discussion on the composition of processing cost in a SSPVT, we refer the readers to the section 1.2 and the section of *The composition of the processing cost* in the general discussion in Tan et al. (2023).

5 Note that all stimuli in our designs are exactly the same, except for the manipulated variables, which are the number of DE operators and their sentential position. Therefore, verification, as such, cannot be a determinant of the effect we measured. We also note that in past studies, the processing/verification issue was addressed directly: in Deschamps et al. (2015), we showed that verification does not interact with the monotonicity effect observed in sentence analysis; in Agmon et al. (2022), we showed that the

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

3.2. *Experimental materials in Hebrew*

When the Hebrew is used, and the NPI *'ey-pa'am* (=ever) replaces *quoi que ce soit*, Hebrew mimics French (cf. Appendix 1 for Hebrew replications of Homer's effects). We test the flip-flop effect in Hebrew with monoclausal sentences, instead of the embedded clausal structures that Homer provided. We use two DE operators, *less* and *not*, to build two double negative structures, (18d) and (18f). These contrast with sentences including zero or one negation, (18a, b, c, e). Note that no NPIs are present, because the experiment is about the processing of multiple negations and domains. Therefore, all sentences are acceptable:

(18) a. [$_{TP\Uparrow}$ [Yoter mi-xezi me-ha-igulim][T [$_{PolP\Uparrow}$ hem kxulim]]]
more than-half of-the-circles are blue
'More than half of the circles are blue' *(0 DE operator; 0 MRD)*

b. [$_{TP\downarrow}$ [**Paxot** mi-xezi me-ha-igulim][T [$_{PolP\Uparrow}$ hem kxulim]]]
less than-half of-the-circles are blue
'**Less** than half of the circles are blue' *(1 DE operator;1 MRD)*

c. [$_{TP\downarrow}$ [**Lo** yoter mi-xezi me-ha-igulim][T [$_{PolP\Uparrow}$ hem kxulim]]]
Not more than-half of-the-circles are blue
'**Not** more than half of the circles are blue' *(1 DE operator; 1 MRD)*

d. [$_{TP\Uparrow}$ [**Lo paxot** mi-xezi me-ha-igulim][T [$_{PolP\Uparrow}$ hem kxulim]]].
Not less than-half of-the-circles are blue
'**Not less** than half of the circles are blue'
(intra-domain structure: 2 DE operators; 0 MRD)

e. [$_{TP\downarrow}$ [Yoter mi-xezi me-ha-igulim][T [$_{PolP\downarrow}$ hem **lo** kxulim]]]
more than-half of-the-circles are not blue
'More than half of the circles are **not** blue' *(1 DE operators; 1 MRD)*

f. [$_{TP\Uparrow}$ [**Paxot** mi-xezi me-ha-igulim][T [$_{PolP\downarrow}$ hem **lo** kxulim]]]
less than-half of-the-circles are not blue
'**Less** than half of the circles are **not** blue'
(cross-domain structure: 2 DE operators; 2 MRDs)

To derive the predictions from the MPH + EnvBA, we compare how many MRDs occurs in each sentence in the relevant minimal pairs in (18). We have three sets of predictions: (i) $RT_{not\ less} < RT_{not\ more}$ because *not less* includes zero MRDs (both PolP and TP are UE), and *not more* includes one (PolP is UE but TP is DE); (ii) $RT_{less\ldots not} > RT_{more\ldots not}$ because MRDs occurs twice in *less*…*not* (the minimal domain PolP reverses from UE to DE, and then TP becomes UE again) but only once in *more*…*not* (an MRD only occurs in PolP); (iii) $RT_{less\ldots not} > RT_{not\ less}$, as *less*…*not* has more MRDs. Note that the prediction (i) has been tested previously in a study which adopted a similar paradigm as the current study ($RT_{not\ less}$ (926.9 ms) < $RT_{not\ more}$ (987.4 ms) and RT_{more} (804 ms) < RT_{less} (888.2 ms), Tan et al. (2023)). By contrast, the prediction of the MPH + OpBA is easy to derive, since

monotonicity effect manipulation of the distance between the sentence that is verified and the time in which the image appears.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

for this hypothesis, the processing cost increments are determined by the number of DE operators.[6]

Before we move on to our experiments, we seek to settle the question of whether the distinct predictions of the OpBA and EnvBA may be confounded with a difference in linear adjacency. That is, whether two DE operators are perceived as a single UE operator by virtue of their immediate proximity rather than being within the same domain. The *flip-flop* effect in (6) and (10) is consistent with this perspective: the NPI *quoi que ce soit* is anti-licensed when the two DE operators are adjacent (6) and (10), but licensed when they are separated (12a-b). To distinguish between domain and linear adjacency, we need an example in which two DE operators are adjacent, but no *flip-flop* effect occurs (or vice versa). Homer shows this for French (19a), and (19b) demonstrates the same effect for Hebrew: (19b) contains two DE operators, *'im* 'if' and *le-xol-ha-yoter xamiša* 'at most five people', and an NPI *ey pa'am* 'ever'.[7] (19a) and (19b) are both grammatical while the two DE operators are adjacent to each other, just like the case in (6). Thus, linear adjacency cannot account for the *flip-flop* effect of (6).

(19) a. **Si au plus cinq personnes** ont fait quoi que ce soit pour aider la Mafia,
If at most five people have done what that this be.SUBJ to help the mafia,
nous sommes sauvés.
we are saved.
'If at most five people did anything to help the Mafia, we are good.'

b. **'im le-xol-ha-yoter** xamiša 'anašim siy'u' ***ey pa'am*** la-mafia, 'anaxnu beseder
if at-most five people assisted ever the Mafia, we good.
'If at most five people ever assisted the Mafia, we are good.'

4. EXPERIMENT I

Experiment I was a lab study, implemented in Hebrew. We used two DE operators in Hebrew, *lo* 'not' and *paxot* 'less', along with *yoter* 'more', to juxtapose between the OpBA and the

6 Two past studies roughly draw a picture similar to ours: Sherman (1973, 1976) found that sentences with the 2∗DE *no one doubted* were easier to process than those with the 1∗DE *doubted*, and were equally difficult to sentences with just *no one* ($RT_{no\ one\ doubted} < RT_{doubted} = RT_{no\ one}$; cf. Schlotterbeck, 2017; Bott et al., 2019, for somewhat related works). Sherman suggested an account that was very much in the spirit of the MPH + EnvBA: he proposed that subjects might mentally combine *no one* and *doubted* to form an affirmative, which are in the same domain. Interestingly, his design contained at least one other condition with these two negatives, namely, sentences containing *doubted...not, which are across two different domains*. And he shows (Table 2, p. 148), that $RT_{no\ one\ doubted} < RT_{doubted....not}$, despite the fact that both had the same monotonicity in the matrix clause. Note that *no one doubted* is an intra-domain 2∗DE structure and *doubted . . . not* is a cross-domain 2∗DE structure. Sherman's results seem to match our prediction regarding the structural difference.

7 The distance between *'im* and *le-xol-ha-yoter xamiša* does not interfere with NPI licensing: even when the two (bolded) DE operators are separated, the sentence is still grammatical, as shown below:
(i) **'im** tagid li še- **le-xol-ha-yoter** xamiša 'anašim siy'u *'ey pa'am* la-mafia, 'ani 'eda
If you-tell me that at-most five people assisted ever the mafia I will-know
še- 'ata mešuga
that you crazy
'If you tell me that at most five people ever assisted the mafia, I will know that you are crazy.'

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

Table 1 Experimental design

		Factor 2: Quantifier Type	
		more	*less*
Factor 1: Sentence Type	**plain**	More than half of the circles are blue.	Less than half of the circles are blue.
	intra-domain *not*	Not more than half of the circles are blue.	Not less than half of the circles are blue.
	cross-domain *not*	More than half of the circles are not blue.	Less than half of the circles are not blue.
	***that*-clause**	It is true that more than half of the circles are blue.	It is true that less than half of the circles are blue.

EnvBA. We asked the participants to verify the aforementioned sentences against pictures in the paradigm of speeded sentence picture verification task. According to the MPH + EnvBA, we would expect to observe $RT_{not\ less} < RT_{not\ more}$ but $RT_{less\ not} > RT_{more\ not}$ based on the number of MRDs. By contrast, if the MPH + OpBA is correct, we would observe $RT_{not\ less} > RT_{not\ more}$ and $RT_{less\ not} > RT_{more\ not}$ instead, since both 2∗DE structures contain one more DE operator than their *more* counterparts.

4.1. *Materials*

We used four Sentence Types along with the pair of polar quantifiers, *more* and *less*, to build up our sentential stimuli. As shown in Table 1, there are eight conditions in total. Sentence Type *plain* was used as a baseline for comparison with other conditions and as a sanity check to make sure the participants are doing the task dutifully. The Sentence Type *intra-domain not* and *cross-domain not*, as discussed in section 2, were the ones which include two DE operators in different configurations. Finally, the Sentence Type *that*-clause was added to counterbalance the number of stimuli with *not* so we would have an equal number of stimuli with and without *not*. The sentences were recorded in Hebrew by a male native speaker, and processed in Audacity to minimize their pitch, amplitude and length variability. However, since the sentences in *that*-clause contain much more words than the other groups, we did not match their length with others (4490 msec vs. 3350 msec), while within group uniformity was controlled.[8]

As for the test images, we used a set of images in which the blue and yellow circles are arranged in a 5 x 5 array. The ratios between the blue and yellow circles are 5:20, 10:15, 15:10 and 20:5. To add variety, for each ratio, there were two types of arrangement of location of the circles. The circles of the same color clustered together to keep the verification task simple.

In order to counterbalance all the conditions, the experiment featured 128 trials (4 Sentence Type x 2 Quantifier Type x 2 Referred Color x 4 Ratio x 2 Arrangement = 128 trials). The Truth Value (true/false) was thus counterbalanced accordingly. Each combination

8 The duration of sentences in *that*-clause was 4490 msec; the duration of sentences in the other 3 conditions was 3350 msec.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

of the trial appeared twice in the experiment. To eliminate the potential learning effect over time, we randomized the order of the trials for each subject.

4.2. *Procedure*

Before beginning an experimental session, the experimenter explained the task to the participant. They were told that they would hear a sentence describing the relationship between circles in two colors, after which, an image with circles in the two colors would appear. Participants were asked to determine whether or not the sentence they heard matched the image, and respond as fast and as accurately as possible. Each experiment included two blocks, each containing 16 practice trials and two runs with 64 trials each. Participants thus responded to 32 practice trials and 256 experimental trials in total. To reduce the difficulty of the task, each block contains only two groups of Sentence Types. Considering the occurrence of *not*, two types of grouping are possible: one is [*plain* + *intra-domain not*] vs. [*that-clause* + *cross-domain not*]; the other is [*plain* + *cross-domain not*] vs. [*that*-clause + *intra-domain not*]. As in the first type of grouping, the combination *plain* and *intra-domain* was the same combination as the stimuli in a previous study (Tan et al., 2023), we decided to assign most of the participants to the second type of grouping and only five participants to the first type of grouping. All the relevant sentences and a few examples of images were shown to participants before the beginning of each block. To minimize fatigue, a 3-minute break was enforced between every two runs.

In every trial, participants first saw a fixation cross on the screen. After 800 ms, they heard a sentence from the headset while the fixation cross remained on the screen. Each sentence was either 3350 ms long or 4490 ms long, depending on condition. Right after the end of the sentence, participants saw an image in the middle of the screen. They were allowed to respond between 300 ms—5000 ms after the onset of the image (Fig. 1).[9] Once participants responded, the trial terminated and the next trial would start. If a participant did not respond within 5000 ms, the response would be recorded as a "miss" and the next trial would start. Participants responded by pressing the right arrow key (TRUE) or the left arrow key (FALSE) on the keyboard. RTs were measured from image onset to key press.

4.3. *Participants*

23 students (12 male and 11 female), aged 26 ± 2.9 (mean ± SD), native Hebrew speakers. All signed an informed consent form approved by the Hebrew University Research Ethics Committee. They received either payment or course credit for participation.

4.4. *Results and analyses*

The data of 22 participants were included (our exclusion criterion was error-based—average accuracy in at least one run < 75%. It led to the removal of only one participant's data). The group's mean accuracy after this screening = 95.12%. In the error domain, within each Sentence Type, the *less* conditions showed lower average accuracy than their *more* counterparts. We show the data of *intra-domain not* and *cross-domain not* in Fig. 2, where

9 In the pilot runs, we started with 1900 ms of response time limit. However, the participants had very poor accuracy in *less...not* condition. Even after we inserted a 500 ms-break between the audio stimulus and the picture, and prolonged the response time limit to 2200 ms, the accuracy of *less...not* condition still remained lower than 70%. In order to curtail the difference in accuracy between different conditions, we eventually decided to prolong the response time limit to 5000 ms.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

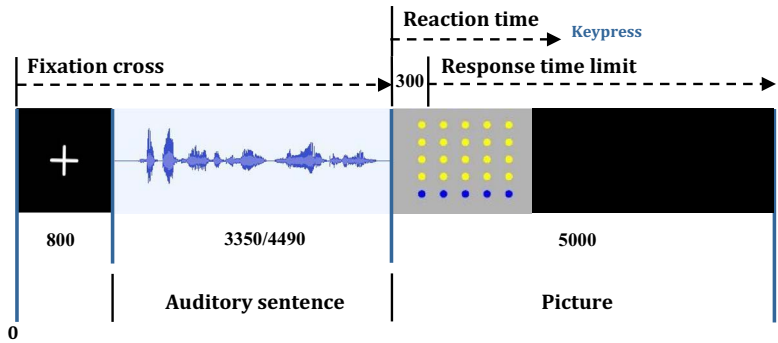

Figure 1 Experiment I timeline (unit: msec)

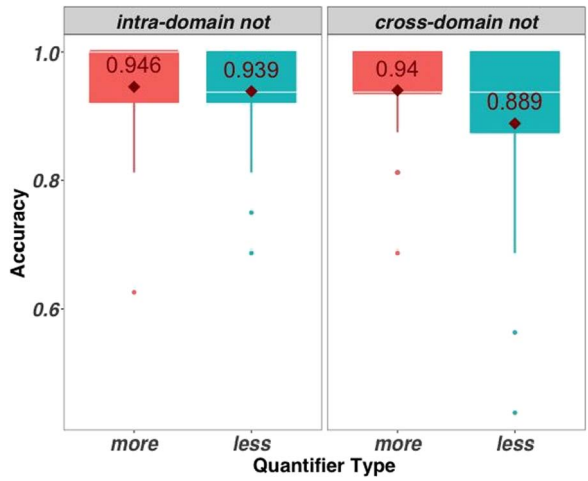

Figure 2 Boxplots of the accuracy data, broken down by Sentence Type and Quantifier Type (coral = *more*; turquoise = *less*). Group distribution statistics are provided as boxes, horizontal midlines and whiskers depicting interquartile range (IQR, defined as first quartile (Q1) to third quartile (Q3)), medians, minimums ($\geq$ Q1–1.5∗IQR), and maximums ($\leq$ Q3 + 1.5∗IQR). In each boxplot, the dark red diamond shows the group mean along with the number. Generally, the accuracy levels were high. Only the accuracy level in the *less...not* condition was slightly lower than 90%.

the mean accuracy of each condition is marked in dark red in each boxplot (the rest of the data: $\text{accuracy}_{more} = 99.23\%$; $\text{accuracy}_{less} = 96.92$).[10] Among all the conditions, the lowest accuracy and widest spread was in *less...not* (88.9%).

When analyzing the RT data, only correct responses were taken into account (since it is unclear what cognitive processes are involved in erroneous responses). The results of the *plain* and *intra-domain not* conditions converge on findings from previous studies (Deschamps et al., 2015; Tan et al., 2023): $\overline{RT}_{more}$ (1106.8 ms) < $\overline{RT}_{less}$ (1389.0 ms) and

10 The Sentence Type that-clause are omitted in the figures and the following analyses because they were for sanity check and for counterbalancing the number of stimuli without *not*, respectively. For those who are curious, the data is close to Sentence Type plain: $\text{accuracy}_{it\ is\ true\ that\ more} = 97.44\%$; $\text{accuracy}_{it\ is\ true\ that\ less} = 95.88\%$; $\text{RT}_{it\ is\ true\ that\ more}$ (1109.5 ms) < $\text{RT}_{it\ is\ true\ that\ less}$ (1364.8 ms).

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

$\overline{RT}_{not\ more}$ (1594.8 ms) > $\overline{RT}_{not\ less}$ (1510.7 ms). As for the pair of conditions named *cross-domain not*, $\overline{RT}_{more..not}$ (1639.9 ms) < $\overline{RT}_{less...not}$ (1885.4 ms). The details of the data of the *intra-domain not* and the *cross-domain not* conditions are in Fig. 3, where means are marked as a dark red diamond. Reviewing the RT data, we notice that: (i) *less . . . not* was the most costly condition in terms of RT; (ii) within every Sentence Type pair, the *less* condition took longer to verify than its *more* counterpart (except the *intra-domain not* condition); (iii) even though both *less . . . not* and *not less* contain two DE operators, the former is much more difficult to process than the latter (longer RT, lower accuracy, and greater variance in accuracy in the former than the latter).

To further test the difference between *not less* and *less . . . not*, we fitted a linear mixed effects model to the data of the *intra-domain not* and *cross-domain not* with the logarithmic transformation of RT as the dependent variable (using R and lme4, Bates et al., 2015). Quantifier Type, Sentence Type and Truth Value and their interactions were used as the fixed effect factors. As random effects, we had an intercept for participants as well as by-participant slope for the effect of Sentence Type.[11] P-values were derived via the lmerTest package in R (Kuznetsova et al., 2017). Visual inspection of residual plots revealed no obvious deviations from homoscedasticity or normality. The statistics are shown in Table 2. Among the statistical results, there are two effects relevant to our research questions: first, there was a main effect of Sentence Type ($t = 4.136$, $p = 0.0004$), as manifested by the fact that the mean RTs of *cross-domain not* were on average higher than those of *intra-domain not*. Secondly, there was a strong significant interaction effect between Quantifier Type and Sentence Type ($t = -6.033$, $p < 0.0001$), indicating that the two double negative structures had very different patterns when compared to their *more* counterpart, respectively.[12]

From the averaged data, we observe that $\overline{RT}_{less...not} > \overline{RT}_{more...not}$, and $\overline{RT}_{not\ less} < \overline{RT}_{not\ more}$. From the results of the mixed effect model, we learned that processing of the pair of *intra-domain not* differs from the processing of the pair of *cross-domain not*, in terms of RT[13].

4.5. *Discussion*

Experiment I tested four Sentence Types with two Comparative Quantifiers in Hebrew, resulting in eight conditions in total, including *more/less* (*plain*), *not more/less* (*intra-domain not*), *more/less...not* (*cross-domain not*) and *it is true that more/less* (*that-clause*). We were particularly interested in the contrast between participants' RTs in the *cross-domain not* condition and those in *intra-domain not*. Even though both *less...not* and *not less* contain two DE operators, the processing difficulty of the two conditions is very different. Our results

11 We chose the model based on the Akaike information criterion (AIC) value. Even though Truth Value did not play a role in our hypothesis, still, including it in the model ameliorates the model fit. Compared to the model which does not contain Truth Value as a fixed effect factor (AIC = 3553.7), the model we adopted in the current analysis has a lower AIC value (= 3502.4), indicating that it was a better fit for the data. Note that the results of the simpler model show the same effect as the more complicated one we used, as shown below.

12 Even in a simpler model: log RT ~ 1 + QType*SentenceType + (1 + SentenceType | Subject), we derived the same effects. There was a main effect of Sentence Type ($t = 4.152$, $p = 0.0004$). Also a significant interaction effect between Quantifier Type and Sentence Type was found ($t = -5.883$, $p < 0.0001$).

13 "contr.sum" was adopted as the contrast scheme in the models for this experiment.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

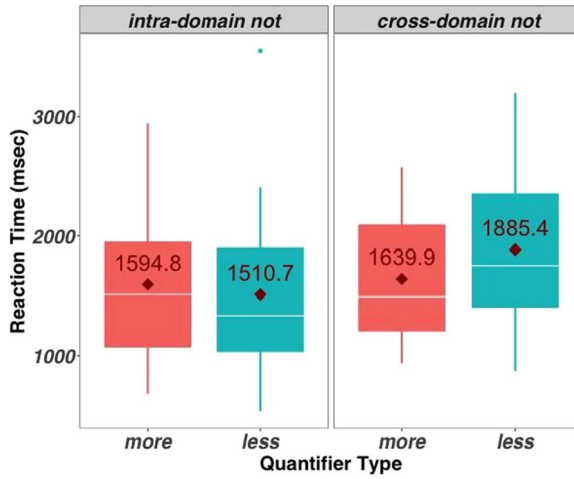

Figure 3 Boxplots of the RT data broken down by Sentence Type and Quantifier Type (coral = *more*; turquoise = *less*). Group distribution statistics are provided as in Fig. 2. Pairwise, *not more* took longer time to process than *not less*; but *more...not* required shorter times to process than *less...not*. In general, subjects took longer to process *cross-domain not* than *intra-domain not*.

Table 2 Summary of the mixed-effect regression model for the data of Experiment I.

	Estimate	Std. error	*t*-value	*P*-value
(Intercept)	7.234	0.07529	96.086	< 0.0001***
QType1	−0.0211	0.00913	−2.312	0.02086*
SentenceType1	0.09102	0.02201	4.136	0.00043***
TruthValue1	0.05639	0.00912	6.186	< 0.0001***
QType1*SentenceType1	−0.05506	0.00913	−6.033	< 0.0001***
QType1*TruthValue1	−0.02155	0.00912	−2.363	0.01818*
SentenceType1*TruthValue1	0.00941	0.00912	1.032	0.30200
QType1*SentenceType1*TruthValue1	0.0346	0.00912	3.796	0.00015***

show that *less...not* had significantly longer RTs and lower accuracy rate than *more...not*, indicating that it was more taxing.[14] By contrast, the *not less* condition yielded slightly shorter RTs and slightly lower accuracy rates than *not more*, and was hence not more difficult to process (perhaps even easier).

The difference between *less . . . not* and *not less* is further corroborated by the results of the linear mixed-effects model when comparing *cross-domain not* and *intra-domain not*, which is also in line with the results of Sherman's study (1976). A related study (Tan et al., 2023) also showed two DE operators in the same domain (i.e., *not less*) facilitate

14 A contrast analysis between $RT_{more\ldots not}$ and $RT_{less\ldots not}$ shows that they are significantly different from each other ($t = 4.828$, $p < 0.0001$). On the other hand, there is no significant difference between $RT_{not\ more}$ and $RT_{not\ less}$ ($t = -2.246$, $p = 0.2169$).

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

the processing of a sentence, suggesting that the two DE operators are counted as a single UE one.

Thus, the processing difference between the *intra-domain* (*not less*) and *cross-domain* (*less...not*) conditions indicates that the processing of two DE operators is not determined by the number of DE operators, suggesting that it is domain-based. In the present context, the notion *domain* is defined syntactically. As shown elsewhere (e.g., Crnič, 2019; Bar-Lev & Fox, 2020), this definition has much semantic and pragmatic relevance. Yet, the nature of this definition leads to the conclusion that syntax is an important determinant of the processing cost of monotonicity. That is, when two DE operators are in the same domain, their processing cost becomes UE (as DE∗DE = UE); but when they are not in the same domain, they cannot be integrated into one during processing and cause MRD, which manifests in elevated cost.

4.6. *Learning effect?*

The account above works, but alternatives must also be considered. One possibility we explored was that the effects obtained were due to selective learning. That is, the differences we observed may have been due to participants' improved performance over the course of the testing session, in a manner that differentiated between the conditions. RTs typically go down in the course of a testing session, and one possibility is that the effect we obtained is due to differential *learning* (some participants reported that they realized that responding is faster if *not less* is converted into *more* during a trial). To see whether this response strategy was indeed used and had a substantial effect on our results, we fitted four regression models for four relevant comparisons, including <*more*, *less*>, <*not more*, *not less*>, <*more . . . not*, *less . . . not* > and <*not less*, *less . . . not*>.[15] We first calculated the cross-participant mean RT_n in trial n ($1 \leq n \leq 32$) in each condition (Fig. 4). Then, we fitted a regression model to each comparison, which included RT as the dependent variable, Trial Order as one predictor, either Sentence Type (for <*not less*, *less . . . not*>) or Quantifier Type (for the other three comparisons) as a second predictor, as well as their respective interaction terms. If differential learning does account for the difference in means within a pair, we expect to observe an interaction effect between Trial Order and one of these pairs. For example, if the result of $RT_{not\ more} > RT_{not\ less}$ is actually a consequence of differential learning, we expect to find an interaction effect in the model for Trial Order and <*not more*, *not less*>. After adjusting for multiple comparisons by Bonferroni corrections (Bonferroni-adjusted p-values = 0.05/4 = 0.0125), we found no significant interaction effects.[16,17] Therefore, learning cannot explain why the pair of *intra-domain not* conditions exhibited a different pattern from the pair of *cross-domain not*, namely, $RT_{not\ more} > RT_{not\ less}$ but $RT_{more\ldots not} < RT_{less\ldots not}$.

15 The other comparisons are not to our interest since they are not minimal pairs, e.g., <*more*, *less . . . not*>, <*not less*, *more . . . not*>, etc.

16 The *p*-values of the interaction terms in each comparison are as below: $p\text{-value}_{<more,\ less>} = 0.0892$; $p\text{-value}_{<not\ more,\ not\ less>} = 0.282$; $p\text{-value}_{<more\ not,\ less\ not>} = 0.323$; $p\text{-value}_{<not\ less,\ less\ not>} = 0.0157$.

17 Even when False Discovery Rate (FDR) is used to correct for multiple comparisons, which is less conservative and more forgiving, no significant interaction effect was found: adjusted $p\text{-value}_{<more,\ less>} = 0.1189$; adjusted $p\text{-value}_{<not\ more,\ not\ less>} = 0.0564$; adjusted $p\text{-value}_{<more\ not,\ less\ not>} = 0.434$; adjusted $p\text{-value}_{<not\ less,\ less\ not>} = 0.0564$.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

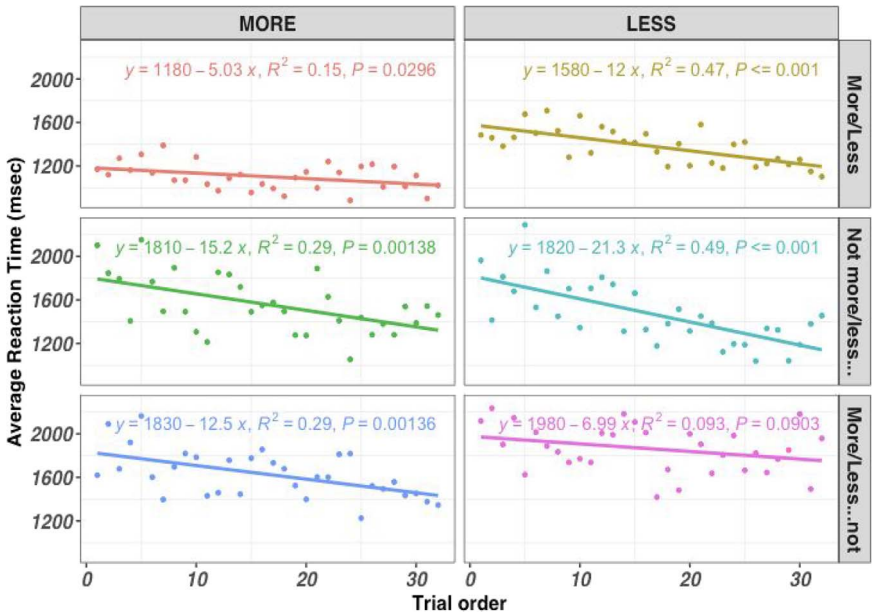

Figure 4 The time course of mean RTs in each condition. The x-axis indicates the order of a trial among all the trials of the same condition. ($1 \leq x \leq 32$, $x \in N$). Only correct responses were taken into account for the mean (86.96% of data). Regression lines were added to show the trend.

5. EXPERIMENT II

Experiment II was a replication attempt of Experiment I, using the exact same conditions and structure, with three differences: (a) participants were recruited on the Internet and the experiment was conducted online; (b) they were native speakers of English, not Hebrew; (c) they constituted a larger group. As will become clear, this experiment further solidifies our results with a large number of participants.

5.1. *Participant recruitment*

We recruited native English speakers online from Prolific's participant pool. Participants were redirected from Prolific.ac to an online experiment hosted on PCIbex Farm. They were promised a monetary reward of £8.00 per hour (which was almost twice the time the test took on average, resulting in an average reward of £4.7). They were also told that completing the study at an overall accuracy of ≥95% would award them a bonus—a completion of their reward to £8.

5.2. *Accuracy-based screening*

1) Upon registration to the study, as well as while signing the consent form, participants were told: "you are required to meet a 90% accuracy threshold. Keep in mind that you may be rejected in the middle of the experiment, due to unsatisfactory accuracy levels. You have 50 minutes to complete the experiment".
2) At the beginning of each run, participants were reminded of the bonus and its requirements.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

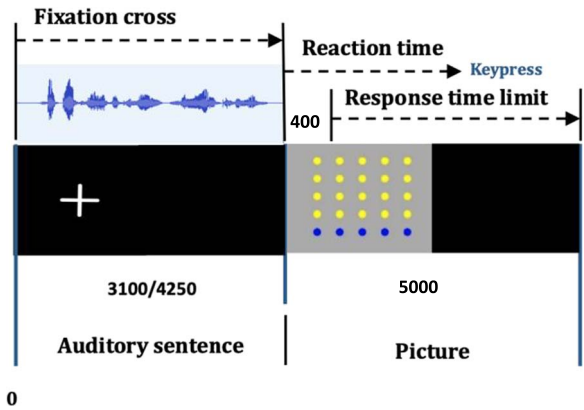

Figure 5 Timeline of Experiment II

3) At the experiment's half point, participants got feedback containing their response accuracy rates, along with a comment on how much closer it brings them to the bonus. Participants whose accuracy rate was < 85% were excluded at this point. Eighty-four participants moved past the half point at the required pace and accuracy, completing the test at an average of 35 minutes. Of these, six exceeded the time limit, or were below the overall 90% accuracy threshold at the experiment's end; they were excluded, but paid. Thirty-three (~ 40%) performed at a level that won them a bonus.

5.3. *Procedures*

After a general explanation about the experiment, each of the two experimental blocks began with a screen containing a graphic representation of a trial's time-course, accompanied by an explanation of the task, and an invitation to perform a single trial. On the next screen, a table with all 8 sentences to be heard in this block were displayed. This was followed by a practice session containing 16 trials (equally representing all conditions) with feedback ("Correct!", "Wrong!", and "Too slow . . . please try to respond faster"). Each trial in both the practice and the experimental sessions was accompanied by a display of the keys representing match and nonmatch. This long preparation phase helped reduce errors, with no bias.

The trial structure was very similar to the one in Experiment I, only with slight time differences in order to adapt to the language differences. Details are depicted in Fig. 5.

After receiving feedback at the experiment's middle point, there was a forced 2-minute break (which the participants could choose to extend to 5 minutes). A forced 1-minute break was also given between the two experimental sessions in each block.

5.4. *Results and analysis*

The data of the 78 participants who completed the experiment were now subjected to more stringent screening criteria: admitted to analysis were only those participants who not only performed at 90% and below 50 minutes, but also, at a level of 75% correct on *each* condition (overall mean accuracy: 95.21%). There were no misses. This screening left us with 70 participants. We show the accuracy data of *intra-domain not* and *cross-domain not*

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

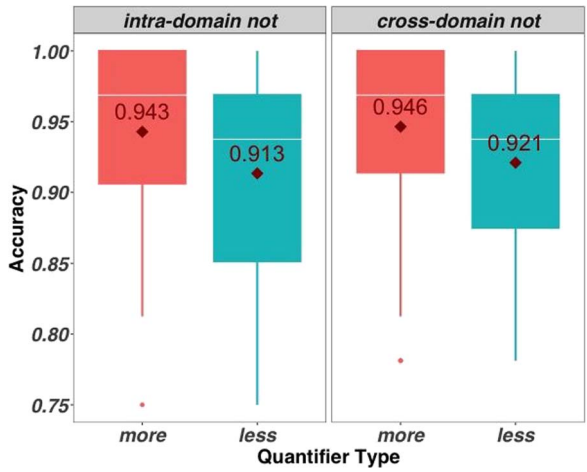

Figure 6 Boxplots of the accuracy data broken down by Sentence Type and Quantifier Type (coral = *more*; and turquoise = *less*). Group distribution statistics are provided in the same way as in Fig. 2. Generally, the accuracy levels were high, all above 90%.

in Fig. 6 (after screening), where the mean accuracy of each condition is marked in dark red in each boxplot.[18] The accuracy level in every condition is above 90%.

Next, we moved from the error domain to the time domain. We analyzed the RTs of the surviving 70 participants, omitting incorrect responses (4.49% of the data). We found that the *less* condition in each Sentence Type took longer to process than its corresponding *more* condition, except in *intra-domain not*, where the RTs of the two Quantifier Types approximate to each other. The mean RTs in *intra-domain not* and *cross-domain not* are exhibited in Fig. 7 (the mean RTs are marked by a dark red diamond in each boxplot.).[19] Despite the difference in experimental method and language, the pattern between the Quantifier Types in each Sentence Type resembles what we found in Experiment I, as shown in Fig. 7. Regarding the two double-DE operator conditions, it seems that *less...not* is much more taxing to process than *not less* because both of the mean and the median of the former are much higher than those of the latter.

To test the significance of the difference between *not less* and *less . . . not*, we fitted a linear mixed effects model to the data of *intra-domain not* and *cross-domain not* with the logarithmic transformation of RT as the dependent variable (using R and lme4, Bates et al., 2015). Quantifier Type, Sentence Type and Truth Value and their interaction were used as fixed effects. The two random-effect terms include an intercept of subjects and a random slope of Sentence Type. P-values were derived via the lmerTest package in R (Kuznetsova et al., 2017). Similar to the result in Experiment I, there was a significant main effect of Sentence Type ($t = -4.719$, $p < 0.0001$) as well as an interaction effect between Quantifier Type and Sentence Type ($t = -11.486$, $p < 0.0001$) (see Table 3). The results here replicate

18 The accuracy of the rest of data is as below: $\text{Accuracy}_{more} = 98.48\%$, $\text{Accuracy}_{less} = 96.79\%$, $\text{Accuracy}_{\textit{It is true that more}} = 98.30\%$ and $\text{Accuracy}_{\textit{It is true that less}} = 95.71\%$.

19 The other mean RTs are: $\text{RT}_{\textit{It is true that less}}$ (1144.8 ms) > $\text{RT}_{\textit{It is true that more}}$ (960.5 ms); RT_{less} (1112.8 s) > RT_{more} (947.7 ms).

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

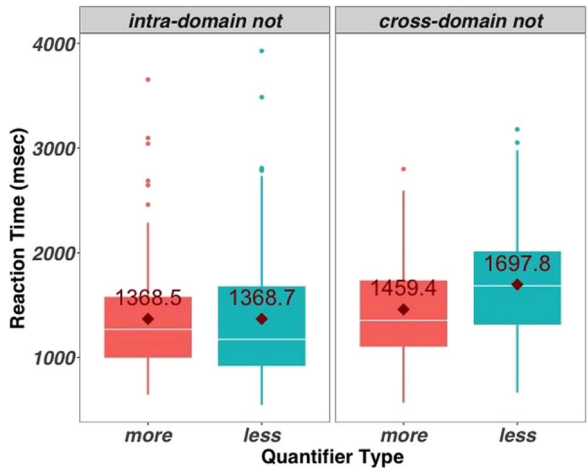

Figure 7 Boxplots of the RT data from the subjects whose accuracies in each condition was over 75%, broken down by Sentence Type and Quantifier Type (the data of 70 subjects are included). Group distribution statistics are provided in the same way as in Fig. 2. The mean $RT_{not\ more}$ is on a par with the mean $RT_{not\ less}$ while the mean $RT_{more...not}$ is smaller than the mean $RT_{less...not}$.

Table 3 Summary of the mixed-effect regression model for the data of Experiment II.

	Estimate	Std. error	*t*-value	*P*-value
(Intercept)	7.140	0.03322	214949	< 0.0001***
QType1	−0.0269	0.00476	−5.648	< 0.0001***
SentenceType1	−0.08468	0.01794	−4.719	< 0.0001***
TruthValue1	0.04046	0.00476	8.498	< 0.0001***
QType1*SentenceType1	0.0547	0.00476	11.486	< 0.0001***
QType1*TruthValue1	−0.00854	0.00476	−1.793	0.07298
SentenceType1*TruthValue1	−0.01516	0.00476	−3.184	0.00146***
QType1*SentenceType1*TruthValue1	−0.02213	0.00476	−4.648	< 0.0001***

the results from Experiment I, indicating that the pattern of *cross-domain not* exhibits a quality different from that of *intra-domain not*.

The online experiment along with the meticulous reward mechanism enabled us to collect a large amount of high-quality data, providing us cross-language evidence which supports *MPH + EnvBA*. We observed the same contrast between *not less* and *less . . . not* among the English speakers as among the Hebrew speakers that *less...not* is much more taxing than *not less*. The fact that we found the same effect in English reinforces our argument that the configuration of a sentence which contains two DE operators determines the way how we process it.

5.5. *Learning effect*

In the same manner as in Experiment I, we calculated the cross-participant mean RT in each trial in each condition for the data in Experiment II, as shown in Fig. 8. Likewise, we fitted four regression models to check the difference between the learning rates in four com-

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

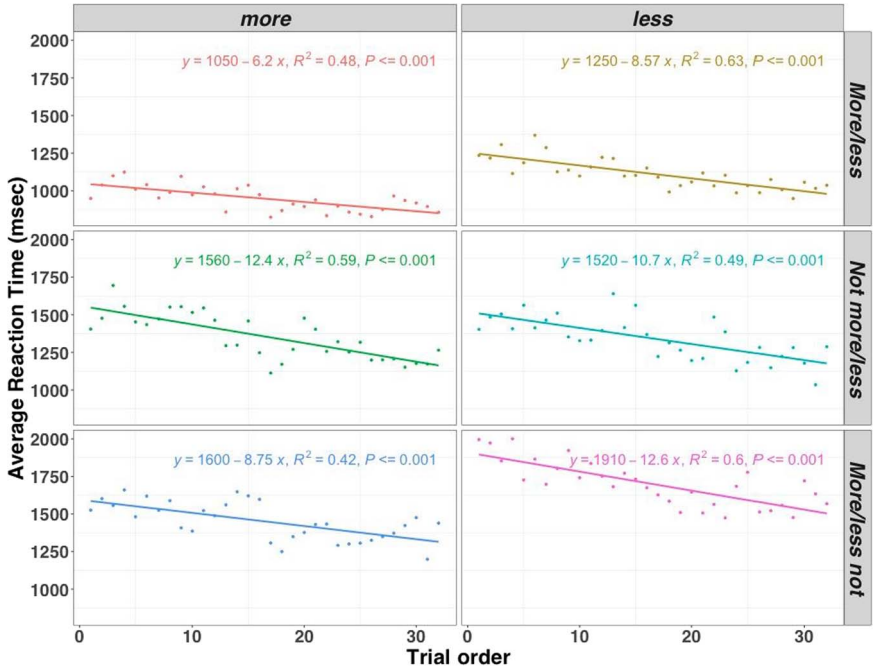

Figure 8 The time course of mean RTs in each condition. The x-axis indicates the order of a trial among all the trials of the same condition. ($1 \leq x \leq 32$, $x \in N$). Only correct responses were taken into account for the mean (95.21% of data). Regression lines were added to show the trend. With the slope of the regression lines, we can compare the learning rates of different conditions.

parisons, including <*more*, *less*>, <*not more*, *not less*>, <*more . . . not*, *less . . . not*> and <*not less*, *less . . . not*>. In each model, RT was the dependent variable, with Trial Order as one predictor, either Sentence Type or Quantifier Type as another predictor, and their interaction term. Among the four comparisons, we found no significant interaction effect, indicating that the learning rates were not distinguishable in each pair.[20] In sum, it was manifested in both Hebrew and English data, that there were no learning rate differences between *more* and *less* conditions in each Sentence Type.

6. GENERAL DISCUSSION

We started by exploring two approaches to NPI licensing, as contrasted by Homer (2010, 2021) in his study of flip-flop phenomena in French. He argued for the EnvBA by showing that only EnvBA could provide a reasonable account of the sentences with two DE operators which do not license NPI. Parallel to Homer's example in French, we showed that there is

20 The *p*-values of the interaction terms in each comparison even before correcting for multiple comparison are all above 0.05, as below: $p\text{-value}_{<more,\ less>} = 0.205$; $p\text{-value}_{<not\ more,\ not\ less>} = 0.573$; $p\text{-value}_{<more\ not,\ less\ not>} = 0.172$; $p\text{-value}_{<not\ less,\ less\ not>} = 0.567$. Hence, after applying FDR corrections, the adjusted *p*-values are even higher, as following: adjusted $p\text{-value}_{<more,\ less>} = 0.410$; adjusted $p\text{-value}_{<not\ more,\ not\ less>} = 0.573$; adjusted $p\text{-value}_{<more\ not,\ less\ not>} = 0.410$; adjusted $p\text{-value}_{<not\ less,\ less\ not>} = 0.573$.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

also a flip-flop of NPI licensing in Hebrew. The Hebrew NPI *ey pa'am* 'ever' is licensed when the two DE operators—*lo* 'not' and *paxot* 'less'—are in different domains, but not when the two sit in the same domain. We then devised the Monotonicity Processing Hypothesis (MPH), which, coupled with EnvBA, predicts that during processing, if domain-wise monotonicity reverses when a lower domain integrates into a higher domain, extra processing cost is induced. In contrast, the MPH + OpBA predicts no effect of syntactic structure on processing. We then provided experimental evidence substantiating the MPH + EnvBA through the measurement of how syntactic structure affects the processing of monotonicity. We demonstrated that, all else being equal, the RT for verifying a sentence containing two DE operators depends on the syntactic relationship between the two. First, compared with their one DE-operator counterparts, respectively, we observed $RT_{not\ less} \leq RT_{not\ more}$ but $RT_{less...not} > RT_{more...not}$ in both Hebrew and English. Secondly, even though the two double DE-operator structures comprise exactly the same words, it took much longer to process the sentence "*less than half of the circles are not blue*" than "*not less than half of the circles are blue*". We suggested that participants were able to integrate two DE operators as one UE operator only when the two DE operators were situated in the same domain.

To conclude, our novel findings—that come from the world of speeded behavior in which the RT, the variable of interest, is continuous—show remarkable convergence with the judgment data that Homer (2021) presented. They therefore provide further evidence to the Environment-based Licensing Condition for NPIs he proposed: NPIs are sensitive to monotonicity of their syntactic environment on the basis of domain because, cognitively, the processing of monotonicity is domain-based. In other words, when two DE operators occur in the same domain, we perceive the domain as one UE domain, resulting in the anti-licensing of NPI in such cases.

Acknowledgements

This research has received funding from an internal grant from Edmond & Lily Safra Center of Brain Sciences, and funding from the European Union's Horizon Europe Programme under the Specific Grant Agreement No. 101147319 (EBRAINS 2.0 Project). This research was supported by the Joint Lab "Supercomputing and Modeling for the Human Brain". We are very grateful to Bernhard Schwarz, Luka Crnič and Emmanuel Chemla for their critical help at different stages of this project. We also thank Jakub Dotlačil and 3 anonymous reviewers for their most helpful comments.

References

Agmon, Galit, Yonatan Loewenstein & Yosef Grodzinsky (2019), 'Measuring the Cognitive Cost of Downward Monotonicity by Controlling for Negative Polarity'. *Glossa: A Journal of General Linguistics* **36**: 1–18.

Agmon, Galit, Yonatan Loewenstein & Yosef Grodzinsky (2022), 'Negative Sentences Exhibit a Sustained Effect in Delayed Verification Tasks'. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **48**: 122–41.

Baker, Carl L. (1970), 'A Note on Scope of Quantifiers and Negation'. *Linguistic Inquiry* **1**: 136–8.

Bar-Lev, Moshe E. & Danny Fox (2020), 'Free Choice, Simplification, and Innocent Inclusion'. *Natural Language Semantics* **28**: 175–223.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker (2015), 'Fitting Linear Mixed-Effects Models Using lme4'. *Journal of Statistical Software* **67**: 1–48.

Bott, Oliver, Fabian Schlotterbeck & Udo Klein (2019), 'Empty-Set Effects in Quantifier Interpretation'. *Journal of Semantics* **36**: 99–163.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

Chierchia, Gennaro (2004), 'Scalar Implicatures, Polarity Phenomena, and the Syntax/Pragmatics Interface'. *Structures and Beyond* **3**: 39–103.

Clark, Herbert H. & William G. Chase (1972), 'On the Process of Comparing Sentences Against Pictures'. *Cognitive Psychology* **3**: 472–517.

Crnič, Luka (2014), 'Non-monotonicity in NPI Licensing'. *Natural Language Semantics* **22**: 169–217.

Crnič, Luka (2019), 'Any: Logic, Likelihood, and Context (Pt. 1)'. *Language and Linguistics Compass* **13**: 1–20.

Deschamps, Isabelle, Galit Agmon, Yonatan Loewenstein & Yosef Grodzinsky (2015), 'The Processing of Polar Quantifiers, and Numerosity Perception'. *Cognition* **143**: 115–28.

Dudschig, Carolin, Ian Grant Mackenzie, Claudia Maienborn, Barbara Kaup & Hartmut Leuthold (2019), 'Negation and the N400: Investigating Temporal Aspects of Negation Integration Using Semantic and World-Knowledge Violations'. *Language, Cognition and Neuroscience* **34**: 309–19.

Fauconnier, Gilles (1975), 'Polarity and the Scale Principle'. *Proceedings of Chicago Linguistic Society* **11**: 188–99.

Gajewski, Jon R. (2005), *Neg-Raising: Presupposition and Polarity*. PhD Thesis,. MIT. Cambridge, MA.

Gajewski, Jon (2011), 'Licensing Strong NPIs'. *Natural Language Semantics* **19**: 109–48.

Gajewski, Jon & I-ta Chris Hsieh (2014), 'Comments on Negative Polarity Items in Definite Description'. *The Art and Craft of Semantics: A festschrift for Irene Heim* 181–198.

Grodzinsky, Yosef, Kim Behrent, Galit Agmon, Nora Bittner, Christiane Jockwitz, Svenja Caspers, Amunts Katrin & Stefan Heim (2021), 'A Linguistic Complexity Pattern That Defies Aging: The Processing of Multiple Negations'. *Journal of Neurolinguistics* **58**: 100982.

Guerzoni, Elena (2006), 'Intervention Effects on NPIs and Feature Movement: Towards a Unified Account of Intervention'. *Natural Language Semantics* **14**: 359–98.

Guerzoni, Elena & Yael Sharvit (2007), 'A Question of Strength: On NPIs in Interrogative Clauses'. *Linguistics and Philosophy* **30**: 361–91.

Homer, Vincent (2010), *Polarity and Modality*. PhD Thesis,. University of California. Los Angeles.

Homer, Vincent (2021), 'Domains of Polarity Items'. *Journal of Semantics* **38**: 1–48.

Just, Marcel A. & Patricia A. Carpenter (1971), 'Comprehension of Negation with Quantification'. *Journal of Verbal Learning and Verbal Behavior* **10**: 244–53.

Klima, Edward (1964), 'Negation in English'. In Jerry A. Fodor and Jerrold J. Katz (eds.), *The Structure of Language*. Prentice Hall. New Jersey. 246–323.

Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen (2017), 'lmerTest Package: Tests in Linear Mixed Effects Models'. *Journal of Statistical Software* **82**: 1–26.

Ladusaw, William A. (1980), 'Affective Or, Factive Verbs, and Negative-Polarity Items'. In J. Kreiman and A. E. Ojeda (eds.), *Papers from the Sixteenth Regional Meeting, III*, pp. 170–184. Chicago: Chicago Linguistics Society.

Papeo, Liuba, Jean-Rémy Hochmann & Lorella Battelli (2016), 'The Default Computation of Negated Meanings'. *Journal of Cognitive Neuroscience* **28**: 1980–6.

Schlotterbeck, Fabian (2017), *From Truth Conditions to Processes: How to Model the Processing Difficulty of Quantified Sentences Based on Semantic Theory*. PhD Thesis. University of Tübingen. Germany.

Schlotterbeck, Fabian, Sonia Ramotowska, Leendert van Maanen & Jakub Szymanik (2020), 'Representational Complexity and Pragmatics Cause the Monotonicity Effect'. *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society* 3398–3404.

Sherman, Mark A. (1973), 'Bound to be Easier? The Negative Prefix and Sentence Comprehension'. *Journal of Verbal Learning and Verbal Behavior* **12**: 76–84.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

Sherman, Mark A. (1976), 'Adjectival Negation and the Comprehension of Multiply Negated Sentences'. *Journal of Verbal Learning and Verbal Behavior* **15**: 143–57.

Tan, I. An, Nitsan Kugler-Etinger & Yosef Grodzinsky (2023), 'Do Two Negatives Make a Positive? Language and Logic in Language Processing'. *Language, Cognition and Neuroscience* **38**: 1027–43.

Tian, Ye & Richard Breheny (2016), 'Dynamic Pragmatic View of Negation Processing'. In P. Larrivée and C. Lee (eds.), *Negation and Polarity: Experimental Perspectives*, pp. 21–43. Cham: Springer.

Tian, Ye, Richard Breheny & Heather J. Ferguson (2010), 'Why We Simulate Negated Information: A Dynamic Pragmatic Account'. *Quarterly Journal of Experimental Psychology* **63**: 2305–12.

Von Fintel, Kai (1999), 'NPI Licensing, Strawson Entailment, and Context Dependency'. *Journal of Semantics* **16**: 97–148.

Wang, Shenshen, Chao Sun, Ye Tian & Richard Breheny (2021), 'Verifying Negative Sentences'. *Journal of Psycholinguistic Research* **50**: 1511–34.

Xiang, Ming, Brian Dillon & Collin Phillips. (2009). 'Illusory Licensing Effects Across Dependency Types: ERP Evidence'. *Brain and Language*, **108**, 1, 40–55.

Xiang, Ming, Julian Grove & Anastasia Giannakidou (2013), 'Dependency Dependent Interference: NPI Interference, Agreement Attraction, and Global Pragmatic Inferences'. *Frontiers in Psychology* **4**: 708.

Xiang, Ming, Julian Grove & Anastasia Giannakidou (2016), 'Semantic and Pragmatic Processes in the Comprehension of Negation: An Event Related Potential Study of Negative Polarity Sensitivity'. *Journal of Neurolinguistics* **38**: 71–88.

Appendix 1: French "flip flop" and our Hebrew experimental materials

This appendix shows the direct connection between the sentence materials of our experiment and the environmental characterization of NPI licensing domains as presented by Homer (2021). Our experiment was conducted in Hebrew main clauses (containing zero, one or two DE operators in subject position). A demonstration that Hebrew is like English and French is thus in order.

We begin with the "flip-flop asymmetry" Homer discusses, inspired by Chierchia (2004), Gajewski (2005) and Guerzoni (2006): a weak NPI is licensed by the DE operator in sentence (A1), but not in (A2). Homer account for this asymmetry by assuming that NPI licensing is syntactically constrained. He posits a Polarity Phrase (PolP) as a domain of NPIs, which DE operators are said to be part of. A weak NPI requires a DE licensor (A1). An NPI is licensed if contained within a PolP with an odd number of DE operators (A2); otherwise, it is anti-licensed (A3).[21]

(A1) $*\ldots[_{TP\Uparrow}$ T $[_{PolP\Uparrow}$ possible $[_{CP}$ que Jean $[_{TP\Uparrow}$T $[_{PolP\Uparrow}$ [*quoi que ce soit*]$_1$ faire $t_1]]]]]$.

(A2) $\ldots[_{TP\downarrow}$ T $[_{PolP\downarrow}$ <u>im</u>possible $[_{CP}$ que Jean $[_{TP\Uparrow}$T $[_{PolP\Uparrow}$ [*quoi que ce soit*]$_1$ faire $t_1]]]]]$.

(A3) $*\ldots[_{TP\Uparrow}$ T $[_{PolP\Uparrow}$ <u>**pas im**</u>possible $[_{CP}$ que Jean $[_{TP\Uparrow}$T $[_{PolP\Uparrow}$ [*quoi que ce soit*]$_1$ faire $t_1]]]]]$.

(A4) $\ldots[_{TP\Uparrow}$ T $[_{PolP\Uparrow}$ <u>im</u>possible $[_{CP}$ que Jean $[_{TP\downarrow}$T $[_{PolP\downarrow}$ <u>**pas**</u> [*quoi que ce soit*]$_1$ faire $t_1]]]]]$.

21 Some examples in Appendix 1 have appeared earlier in the main text. They are iterated here for the readers' convenience.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

Example (A5), where DE-ness is induced in the antecedent clause of a conditional sentence, confirms Homer's hypothesis. In (A5) a PolP is found, that contains the NPI, as well as a single DE operator, ***im**possible*. Thus, the second DE operator in the antecedent of the conditional, ***si***, is above the licensing domain, unlike in (A3), where both are within this domain:

(A5) [$_{TP\Uparrow}$ [$_{CP\Uparrow}$ **Si** … **T** [$_{\mathbf{PolP}\Downarrow}$ **im**possible [$_{CP}$ que Jean [T [$_{PolP\Uparrow}$ [*quoi que ce soit*]$_1$ faire t$_1$]]]]] T] …

In addition to PolP, Homer posits TP as an another domain of NPIs, based on (A6), which contains two DE operators, *si* "if" and *au plus cinq personnes* "at most five people". Assuming the subject *au plus cinq personnes* sits at [Spec, TP], for *quoi que ce soit* to be licensed, there must be a domain between CP and PolP, namely, TP, as Homer proposes.

(A6) [$_{TP\Uparrow}$ [$_{CP\Uparrow}$ **Si** [$_{TP\Downarrow}$ [**au plus cinq** personnes] T [$_{PolP\Uparrow}$ ont [*quoi que ce soit*]$_1$ faire t$_1$]]] T] …

Hebrew offers a perfect replication of the examples that get Homer's theory off the ground (with the weak NPI *'ey-pa'am* = ever):

(A1-heb) * … [$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ efšari [$_{CP}$ še Dani [$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ nirdam *'ey-pa'am*
possible that Dani fell.asleep ever
be-šmira]]]]]
while.on.guard
'… possible that Dani ever fell asleep while on guard'.

(A2-heb) … [$_{TP\Downarrow}$ T [$_{PolP\Downarrow}$ **bilti**-efšari [$_{CP}$ še Dani [$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ nirdam *'ey-pa'am*
im-possible that Dani fell.asleep ever
be-šmira]]]]]
while.on.guard
'… impossible that Dani ever fell asleep while on guard'

(A3-heb) * … [$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ **lo bilti**-efšari [$_{CP}$ še Dani [$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ nirdam *'ey-pa'am.*
not im-possible that Dani fell.asleep ever
be-šmira]]]]].
while.on.guard.
'… not impossible that Dani ever fell asleep while on guard'

(A4-heb) … [$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ **bilti**-efšari [$_{CP}$ še Dani [$_{TP\Downarrow}$ T [$_{PolP\Downarrow}$ **lo** nirdam *'ey-pa'am.*
im-possible that Dani not fell.asleep ever
be-šmira]]]]].
while.on.guard.
'… impossible that Dani did not ever fall asleep while on guard'

(A5-heb) … [$_{TP\Uparrow}$ [$_{CP\Uparrow}$ **'im** [$_{PolP\Downarrow}$ **bilti**-efšari [$_{CP}$ še Dani [$_{TP\Uparrow}$ T [$_{PolP\Uparrow}$ nirdam *'ey-pa'am.*
if im-possible that Dani fell.asleep ever
be-šmira]]]]] T] …
while.on.guard.
'If it is impossible that Dani ever fell asleep while on guard, …'

(A6-heb) … [$_{TP\Uparrow}$ [$_{CP\Uparrow}$ **'im** [$_{TP\Downarrow}$ [**le-xol-ha-yoter** xamiša 'anašim] T [$_{PolP\Uparrow}$ siy'u
if at.most five people assisted
'ey-pa'am la mafia]]] T] …
ever to.the. mafia.
'If at most five people ever assisted the mafia, …'

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024

Appendix 2: extension of "flip flop" to simple, monoclausal sentences

Embedded clauses and conditionals turn out to not be the only flip-flop environments. Homer presents (A7), a monoclausal sentence containing two DE operators, *au plus cinq* and *n'…pas*. Again, assuming the subject *au plus cinq* is at [Spec, TP], there must be a DE domain below TP, which licenses the NPI, namely, PolP.

(A7) *Au plus cinq* personnes n' ont pas fait quoi que ce soit
at most five people ne have neg done what that this be.SUBJ
pour aider la Mafia.
to help the Mafia.
'At most five people didn't do anything to help the Mafia.'

Unfortunately, Homer does not provide a minimal contrast where *at most* combines with a constituent negation. This is important to our experiment, as our materials were simple double-negative sentences with either constituent negation or sentential negation. Below we show that Homer's analysis holds for simple sentences, as a "flip-flop asymmetry" is found elsewhere.

We used mono-clausal sentences whose subjects are generalized quantifiers with modified numerals (*yoter/paxot me-xamiša* = more/less than five), and their negated counterparts (*lo yoter/paxot me-xamiša* = not more/less than five). A weak NPI (*'ey pa'am* = ever) in object position is not licensed by a UE-quantifier (A8); yet it is licensed by the DE-quantifier *paxot* (A9). As in (A7), when a sentential negation *lo* is added to a sentence with a DE operator in the subject position, an NPI is still licensed (A10). The flip-flop happens when a constituent negation is added to (A9), as no DE domains are available in this case (A10).

(A8) *[$_{TP\Uparrow}$ [Yoter me- xamiša ratzim][T [$_{PolP\Uparrow}$ [higi'u *'ey pa'am* la-gmar]]]].
more than-five runners reached ever to.the-finish.linet
'More than five runners have ever passed the finish line.'

(A9) [$_{TP\downarrow}$ [**paxot** me-xamiša ratzim][T [$_{PolP\Uparrow}$ [higi'u *'ey pa'am* la-gmar]]]].
less than-five runners reached ever to.the-finish.linet
'Less than five runners ever reached the-finish-line.'

(A10) [$_{TP\Uparrow}$ [**paxot** me-xamiša ratzim] [T [$_{PolP\downarrow}$ **lo** [higi'u *'ey pa'am* la-gmar]]]].
less than-five runners not reached ever to.the-finish.line.

(A11) *[$_{TP\Uparrow}$ [**lo paxot** me-xamiša ratzim] [T [$_{PolP\Uparrow}$ [higi'u *'ey pa'am* la-gmar]]]].
not less than-five runners reached ever to.the-finish.line

These examples indicate that Homer's effects can be extended to monoclausal sentences. The two DE operators in (A11) integrate and together provide a UE environment to *ey pa'am*, whereas in (A10), the *not* itself may provide a DE domain to *ey pa'am*. The key point here is how the integration of two DE operators occurs, since eventually, in the maximal domain, i.e. the sentence, *ey pa'am* is in a UE domain in both (A10) and (A11). The contrast between (A10) and (A11) suggest that the integration happens stage by stage (*cf*. Papeo et al., 2016).

These Hebrew data align with Homer's hypothesis that PolP and TP are both valid domains for NPIs, showing that our experimental materials are well-suited to distinguish between the hypotheses.

Downloaded from https://academic.oup.com/jos/article/41/1/77/7656700 by Forschungszentrum Julich , Zentralbibliothek user on 03 July 2024