Generalizing Polarity Sensitive Items to Positive Environments in an Artificial Language

Jeremy Kuhn[1] and Mora Maldonado[2]

[1] Institut Jean Nicod, CNRS, ENS, EHESS, PSL University
jeremy.kuhn@ens.psl.eu
[2] LLING, CNRS, Université de Nantes
mora.maldonado@univ-nantes.fr

Abstract

Two well known patterns involve polarity sensitive items. First is Jespersen's cycle: a polarity sensitive item appearing in the scope of negation is reanalyzed as itself the marker of negation. Second, many NPIs in English have a secondary positive use, with a systematic semantic relation to the polarity sensitive use: in each case, the meaning of the positive form is the logical dual of the NPI. This study investigates these meaning changes using an artificial language learning paradigm, where learners are trained on ambiguous input, and must generalize beyond this input. We ask which generalization strategies appear when learners have to interpret polarity sensitive items in a positive environment.

1 Introduction

Two well known patterns involve polarity sensitive items. The first is a diachronic pattern known as Jespersen's cycle [3, 5]: a polarity sensitive item appearing in the scope of negation is often (historically) reanalyzed as itself the marker of negation, as shown for French in (1).

(1) Jespersen's cycle
a. Je ne comprends. → b. Je ne comprends pas. → c. Je comprends pas.

The second pattern regards Negative Polarity Iitems (NPIs), which often have a secondary positive use [6], as shown for English in (2). Ladusaw observes that the semantic relation between the two uses is systematic: in each case, the meaning of the positive form is the logical dual, the meaning one would have to assign if the NPI were scoping over negation, as illustrated in (3). Other languages have similar patterns, e.g. French *encore*, 'yet/still'.

(2) Dual uses of NPIs
a. I talked to any person who'd listen. (*any* = *every*)
b. I've been here ever since you called. (*ever* = *always*)
c. The winner is yet to be decided. (*yet* = *still*)
d. Gas is expensive anymore. (*anymore* = *nowadays*)

(3) a. not any = every not
b. not ever = always not
c. not yet = still not
d. not anymore = nowadays not

1.1 A diachronic view

These meaning shifts can be understood from a diachronic perspective. Let us start by considering a child learning a language with polarity sensitive items. In principle, the child should be

able to correctly identify the meaning of the sentence as a whole. However, due to distributional constraints (i.e., the fact that some items consistently co-occur together), multiple equally viable analyses are possible for the meaning of the individual words and the way they are put together. If the syntactic constraints are relaxed, as it may happen in language evolution, these different analyses would lead to observably different sentential meanings.

For Jespersen's cycle, consider a snapshot of a language at stage (1)b. A child learning language will successfully understand that the sentence as a whole has a negative meaning (in this case, 'the speaker doesn't understand'), and can deduce the meanings of *je* and *comprends* from their use in other sentences. But, because *ne* and *pas* always occur together, the child does not know which element contributes the negation—there are two unresolved variables. Postulating that *pas* has negative meaning allows the development of stage (1)c, where the syntactic constraint of co-occurrence has been removed.

Dual uses of NPIs can be viewed through a similar diachronic lens. Consider a snapshot of a language in which a particular lexical item always occurs in the presence of negation. In such a situation, two equally viable analyses are possible to derive the meaning of a sentence containing the item: either it has one meaning (e.g., $\exists$) and scopes below negation, or it has another meaning (e.g., $\forall$) and scopes above negation. Once again, there are two unresolved variables: the denotation of the lexical item, and its scope (above or below negation).

If the syntactic constraint is removed (e.g., if an NPI appears in a non-negative context), this should reveal the (preferred) underlying interpretation of the polarity sensitive item as well as its relative scope. The examples in (2) thus suggest a stage where many NPIs are interpreted as wide-scoping elements with the appropriate (dual) meaning. Note that this is not incompatible with the existence of another (earlier or later) stage with the same items are interpreted as narrow-scoping.

For both processes described above, when a particular lexical item is restricted to a specific syntactic environment, there may be a systematic ambiguity of analysis that, in its resolution, ends up acting as a vector of grammatical change.

1.2 Current study

The present study investigates these possible meaning changes using an artificial language learning paradigm, where learners are trained on ambiguous input, and must generalize beyond this input in a way that is compatible with only one of the hypotheses of interest (a.k.a. 'Poverty of stimuli' paradigm [2, *i.a.*]). Here, we ask which generalization strategies appear when learners have to interpret polarity sensitive items in a positive environment.

For sentences of the form [... Neg ... X ...], where X only appears in the presence of a word independently learned to denote sentential negation, three possible generalization strategies are possible when X appears on its own:

(4) (i) a narrow-scope (often existential) meaning (=the 'literal' NPI meaning);
(ii) a wide-scope (often universal) meaning (=its logical dual); or
(iii) a redundant negative meaning, analogous to Neg itself (=Jespersen's cycle).

In Experiment 1 (Exp 1), English-speaking participants were taught an artificial language that included a negative marker and a polarity sensitive degree modifier roughly equivalent to English *at all*. All participants were trained on sentences where this item always co-occurred with negation. That is, the input did not include any example of this item in absence of negation. At test, learners were asked to assign an interpretation to sentences where this novel item appeared on its own, without negation. In Experiment 1a (Exp 1a), we contrasted strategies (i)/(iii) and (ii); in Experiment 1b (Exp 1b), strategies (i)/(ii) and (iii). Experiment

2 (Exp 2) repeated Exp 1a but with a polarity sensitive item quantifying over times, akin to English *ever*.

All experimental materials, data, and analysis reported here are available at osf.io/kzju8. Experimental design and analysis were preregistred for Exp1a (Exp 1b is identical) and Exp2.

2 Experiment 1

2.1 Methods

Design and Materials The miniature language consisted of four predicates, four proper nouns, one negative marker (*em*) and one degree modifier (*tup*), akin to English *at all*. All words were created following English phonotactics and presented orthographically. Predicates denote gradable properties with closed scales: transparent/opaque, open/closed, top/bottom, and standing/lying. For each property, we define four possible scale points (or degrees): 'minimum', 'near-minimum', 'near-maximum', and 'maximum', as shown in Table 1 for 'transparent'. The predicate-meaning mapping was manipulated between subjects such that the maximum scale point could be linked to either end of the scale (e.g., the maximum degree could be 'opaque' or 'transparent').

Participants were exposed to non-target and to critical sentences. Non-target sentences were: *simple positive (SP)* (e.g., *Greenie pleet*), *simple negative (SN)* (e.g., *Greenie em pleet*), and *negative 'at all' (Neg-'At all')* (e.g., *Greenie em pleet tup*). The meaning of these sentences was conveyed by showing participants whether they could be used to describe scenarios where the predicate applied to the different degrees, as in Table 2a.

The critical configuration, seen only during testing, was the *positive 'at all' (Pos-'At all')* sentence, where the modifier *tup* appeared on its own, without negation (e.g., *Greenie pleet tup*). In Exp 1a, participants had to decide whether Pos-'At all' sentences described a scenario where the predicate applied to a maximal or to a near-minimal degree. In Exp 1b, participants were asked to choose whether the predicate applied to a maximal or to a minimal degree.

Exp 1a and 1b together allow us to dissociate the generalization patterns in (4): If learners interpret *tup* as an universal quantifier, they are expected to go for the 'maximum degree' interpretation across the board. Similarly, if *tup* is treated as a redundant negative marker, both 'near-minimum' and 'minimum degree' interpretations would be compatible with the sentence. In contrast, if *tup* is treated as an existential quantifier, learners are expected to go for the 'near-minimum degree' reading in Exp 1a but for the 'maximum degree' one in Exp 1b.

Procedure The experimental procedure was the same in Exps 1a and 1b. Participants were instructed that they were going to learn a foreign language called Hiri Motu through a training and a testing phase. In the training phase, participants were taught how to interpret non-target sentences. The training consisted of (a) an exposure block, where participants were presented with scenarios that verify the sentence, and (b) a sentence-picture matching block, where participants were presented with a sentence together with two pictures and had to select which picture best described the sentence (see Fig. 1A). Participants were given feedback on their answers. There were 14 exposure trials and 24 sentence-picture matching trials.

The testing phase involved a similar matching task but also included trials for the critical, held-out Pos-'At all' sentence. The alternative pictures for this held-out sentence were determined by the experiment (see Fig. 1B). There were 32 testing trials (8 involving the critical sentence configuration). No feedback was provided to participants at this stage.

At the end of the experiment, participants were asked to translate to English four sentences from the artificial language, one of each type.

Participants 132 English-speaking adult participants were recruited via Prolific (Exp 1a: 75, Exp 1b: 57). Participants were paid 9£/hour for their participation which lasted on average 15 minutes. Per our pre-registration, we only included in the analysis participants who were successfully trained on non-target sentences (i.e. accuracy rates above 75%). This resulted on the analysis of 49 participants in Exp 1a and 33 in Exp 1b.

2.2 Results and discussion

Fig. 3a shows the proportion of trials on which participants chose the 'maximum degree' scenario for Pos-'At all' sentences during the test phase. A visual inspection suggests that in both experiments about a quarter of participants interpreted Pos-'At all' sentences as compatible with scenarios where the predicate applies to maximum degree, but half the participants generalized to the other choice, whatever it was. Per our pre-registration, we evaluated this pattern statistically with logistic mixed-effects regression models, including random intercepts per subject.[1] These models revealed that the proportion of 'maximum-degree' responses was significantly below chance in Exp 1a ($\beta = -1.563; p = .0137$) and not different from chance in Exp 1b ($\beta = -0.89; p = .2$).[2]

In order to compare the two experiments, we fit a third logistic regression, predicting responses by Experiment. The effect of Experiment was not significant ($\beta = .69$; $p = .46$). This suggests that participants who interpreted Pos-'At all' sentences as compatible with 'near-minimum degree' scenarios in Exp 1a learned something analogous to a redundant negative meaning, and not a proper existential meaning.

Translations provided at the end of Exp 1a give further support for this interpretation. 'Maximum degree' responders systematically translated Pos-'At all' sentences with universal degree modifiers like *completely* or *very* and Neg-'At all' sentences as involving either the same words (e.g. *Greenie is very transparent*) or the modifier *at all* (e.g. *Greenie is not transparent at all*). Among the 'near-minimum degree' responders, translations of Pos-'At all' sentences were less consistent; no subject translated the meaning using existential degree modifiers like *a bit*, but a number of participants gave the sentence a *negative* meaning.

Altogether, Exp 1 results suggest that learners tend to interpret the degree modifier *tup* as an additional, redundant negation. This modifier can co-appear with negation, as in Neg-'At all' sentences, giving rise to an emphatic negative meaning, or appear on its own, leading to the same meaning as the alternative negative marker *em*.

Now, polarity sensitivity is known to vary within and across languages, possibly as a function of the lexical semantics of the items themselves [4]. In Exp 1, 'at all' was chosen as the target meaning because English *at all* is one NPI that that does not have a positive dual use. On the other hand, it is possible that lexical semantic properties of *at all* also allow it to be more easily interpreted as a bleached marker of sentential negation. In Exp 2, we thus modified the experimental materials so the artificial language involves a different polarity-sensitive item: a temporal adverb roughly equivalent to the English word *ever*.

[1] All statistical analyses were performed using the lme4 package in R [7, 1]. In the models reported here, random intercepts per predicate and nesting by group had to be removed due to lack of convergence.

[2] The low number of participants in Exp 1b might have resulted on a lack of sensitivity to the effect. In an initial stage, the exclusion threshold for Exp 1b was of 66% accuracy rates in non-target trials (rather than 75%); only later we noticed that this was not in accordance with what was preregistered for Exp 1a. Analysis of the data with the lower exclusion threshold showed a statistical effect in the same direction as Exp 1a.

3 Experiment 2

3.1 Methods and materials

The methodology in Exp 2 was analogous to Exp 1a, except in the details specified below.

Design and Materials Participants were presented with a miniature language which in this case included one copula, three predicates, one proper noun, one negative marker (*mo*) and one adverb (*yele*), akin to English *ever*. Predicates denote three possible locations in a map: city, desert, forest. All sentences in the artificial language refer to either the current location or the past trajectory of a character, Jojo. For example, a sentence can express that Jojo is currently in the city, or that they have never been in the city. For each location, we define four possible temporal states of Jojo: 'never', 'previously', 'currently', and 'always', as shown in Table 3 for the city location. Note that one can establish a correspondence between Jojo's temporal states for each location, and the scale points or degrees used in Exps 1a and 1b.

As in Exp 1, participants were first trained on three non-target sentences: (i) *simple positive (SP)* (e.g., *Jojo es villum*), used either when Jojo has always been or currently is in the location denoted by the predicate; (i) *simple negative (SN)* (e.g., *Jojo mo es villum*), used either when Jojo has previously been or has never been in the location denoted by the predicate; and (iii) *negative 'Ever' (Not-'Ever')* (e.g., *Jojo mo es villum yele*), used exclusively when Jojo has never been in the location denoted by the predicate.

At test, participants were asked to additionally provide a meaning for *positive 'ever' (Pos-'Ever')* sentences (e.g., *Jojo es villum yele*), where *yele* does not co-appear with negation. Participants had to decide whether this critical sentence means that Jojo has always been in the given location or whether they have previously been in this location (but no longer are). As in Exp 1a, this choice serves to distinguish between strategies (i)/(iii) vs. (ii) in (4).

Procedure The procedure was analogous to Exp 1. There was a training and a testing phase, including 22 exposure trials, 32 training trials and 32 testing trials (8 involving the Pos-'Ever' sentences). A trial example is provided in Fig. 2.

Participants We recruited 59 English-speaking participants via Prolific. 31 participants were successfully trained on non-target sentences and taken into account for the analysis.[3]

3.2 Results and discussion

Fig. 3b shows the proportion of trials on which participants chose the 'always' scenario for Pos-'Ever' sentences during testing. Per our pre-registration, we fit a logistic mixed effects regression (random intercepts per subject) to statistically evaluate these responses. Our model reveals that the log-odds of selecting the 'always' scenario is significantly above chance ($\beta = 9.5; p < .001$). This result contrast with Exp 1, as a majority of participants appear to consistently take the polarity item to have a universal meaning in positive environments, consistent with a logical dual interpretation (i.e., strategy (ii) in (4)).

4 General Discussion

Our findings reveal that different strategies are adopted when extending the meaning of polarity sensitive items to contexts without a licensor. These strategies correspond to two meaning shifts

[3]Training items turned out to be difficult to learn, and for this reason we have not been able to collect our target sample size in this experiment.

attested cross-linguistically: (i) Jespersen's cycle and (ii) positive wide-scope duals. In Exp 1a, most participants assign a 'near-minimum degree' interpretation to Pos-'At all' sentences. The comparison with Exp 1b suggests that this is not due to an existential interpretation, but rather to a negative interpretation of the NPI, possibly due to a repair strategy with reconstructed negation. This generalization corresponds to Jespersen's cycle, in which a minimizing NPI is reinterpreted as contributing negation itself.

In Exps 1a and 1b, we additionally observe a small group of participants who assign a 'maximum degree' interpretation to Pos-'At all' sentences, thus displaying a pattern of generalizing the NPI meaning to its wide-scope dual. Exp 2 corroborates the existence of this group, and suggests that strategies differ depending on the semantic properties of the polarity sensitive item in question: *at all* is a better marker of redundant negation than *ever*.

Differences between NPIs Exps 1 and 2 displayed different generalization preferences. Since the two tasks were not perfectly parallel, it is possible that these differences arose in part due to irrelevant aspects of the experimental design. On the other hand, it is also possible the different generalization preferences arise from differences in the lexical semantics of the two NPIs. Why might *at all* be a better marker of redundant negation than *ever*?

One possible explanation concerns the relation of the NPIs to a certain kind of context-sensitivity. Specifically, simple negative sentences show vagueness with respect to degree quantification that they do not show with respect to time quantification. Consider the sentence in (5) with the two scenarios described in (6). While the sentence could easily be judged to be true in either scenario, one can nevertheless judge Scenario (b) as a 'better' instantiation of the sentence. In contrast, consider the same sentence with the two scenarios in (7). Again, the sentence is true in either scenario. Here, though, there is no clear intuition that either scenario is 'better' than the other.

(5) The gas tank is not full.

(6) a. The tank is a quarter full.
b. The tank is completely empty.

(7) a. The tank is completely empty, but was full yesterday.
b. The tank is completely empty, and has been all week.

The difference in intuitions can be explained in terms of context-sensitivity. Notably, (5) has a context-sensitive degree variable that may take different values depending on the question under discussion. (One needs to know *how full* counts as 'full' to evaluate the sentence.) In contrast, the time variable in (5) shows no context-sensitivity, as it is saturated by the time of utterance. (One doesn't need to know *how often* the tank is full to evaluate the sentence.)

As a consequence, the meaning of *at all* can more easily be bleached than the meaning of *ever*. In certain contexts, the threshhold of *full* is sufficiently low that (5) is synonymous with *The tank is not at all full*. In contrast, *ever* introduces temporal quantification not present in (5); there is thus no context in which (5) is synonymous with *The tank is not ever full*. The inability to ignore the temporal component of *ever* may thus explain why it resists the bleached negative meaning that appears as the dominant generalization strategy in Experiment 1.

The influence of context-sensitivity, as well as other aspects of the lexical semantics, can be investigated in follow-up experimental work with other polarity sensitive items ('yet', 'anymore'), as well as typological work on sycretisms involving NPIs cross-linguistically.

References

[1] Douglas M Bates. lme4: Mixed-effects modeling with r, 2010.

[2] Jennifer Culbertson and David Adger. Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences*, 111(16):5842–5847, 2014.

[3] Östen Dahl. Typology of sentence negation. *Linguistics*, 17(1-2):79–106, 1979.

[4] Michael Israel. Polarity sensitivity as lexical semantics. *Linguistics and philosophy*, pages 619–666, 1996.

[5] Otto Jespersen. *Negation in English and other languages*. Høst, Copenhagen, 1917.

[6] William Ladusaw. *Polarity sensitivity as inherent scope relations*. PhD thesis, University of Texas at Austin, 1979.

[7] R Core Team. R: A language and environment for statistical computing, 2018.

minimum	near-minimum	near-maximum	maximum
			

Table 1: Scenarios in Exp 1. Illustration of four scale points (degrees) for the gradable property 'transparent', which is expressed by the predicate *pleet* in the experiment.

(a) Experiment 1 (see Fig. 1)

Scenario	SP	SN	Neg-NPI
Minimum degree	✗	✓	✓
Near-minimum degree	✗	✓	✗
Near-maximum degree	✓	✗	✗
Maximum degree	✓	✗	✗

(b) Experiment 2 (see Fig. 3)

Scenario	SP	SN	Neg-Ever
Never	✗	✓	✓
Previously	✗	✓	✗
Currently	✓	✗	✗
Always	✓	✗	✗

Table 2: Scenarios that verify non-target sentences in each experiment. The ✓sign indicates that the sentence could be used in the scenario; ✗ that it could not.

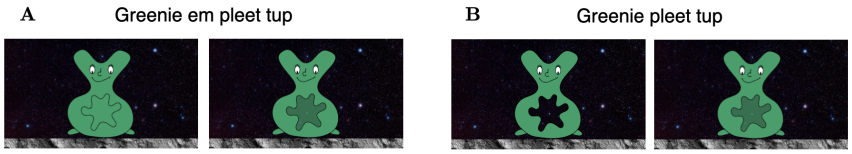

Figure 1: Sentence-picture matching trials in **(A)** Training (Neg-'At all') and **(B)** Testing (Pos-'At all') in Exp 1a.

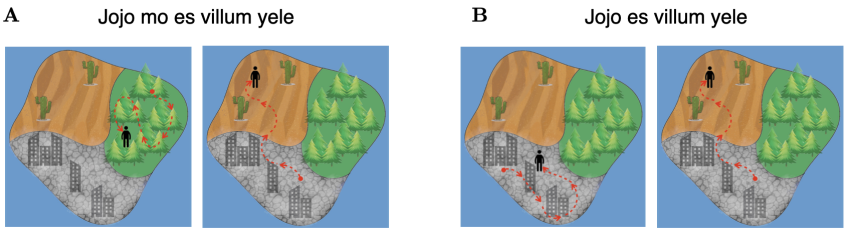

Figure 2: Sentence-picture matching trials in **(A)** Training (Neg-'Ever') and **(B)** Testing (Pos-'Ever') in Exp 2.

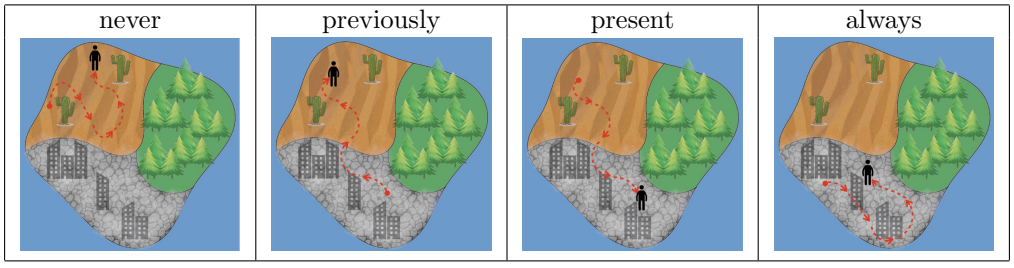

Table 3: Scenarios in Exp 2. Illustration of the four possible states for the city location, which is expressed by the predicate *villum* in the experiment.

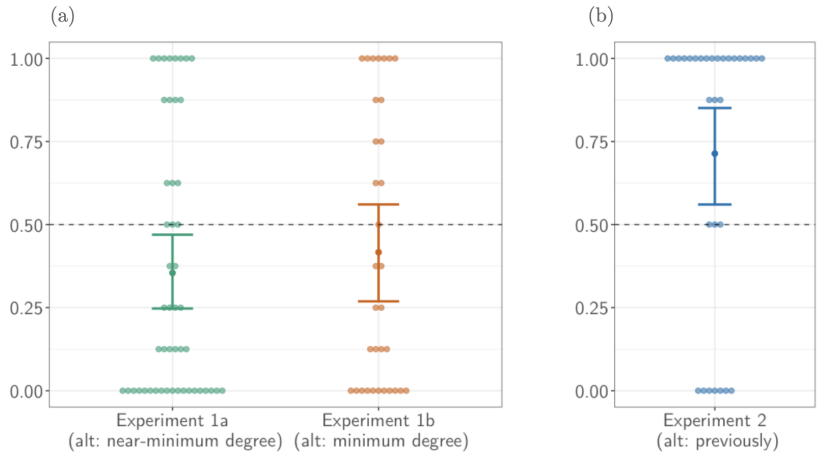

Figure 3: Results in Exps 1 and 2. (a) Proportion of 'maximum degree' responses for Pos-'At all' sentences in Exps 1a and 1b. (b) Proportion of 'always' responses for Pos-'Ever' sentences in Exp 2. Dots are individual participants, error bars are 95% bootstrapped confidence intervals.