

Polarity Judgments: An Empirical View

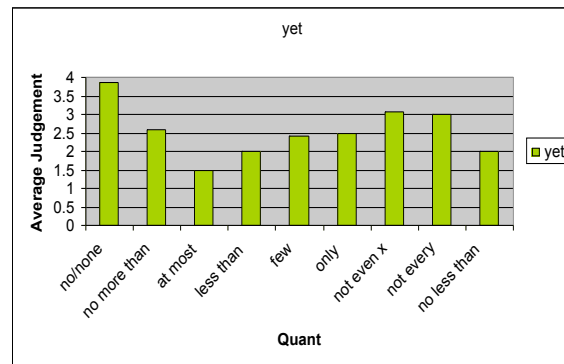
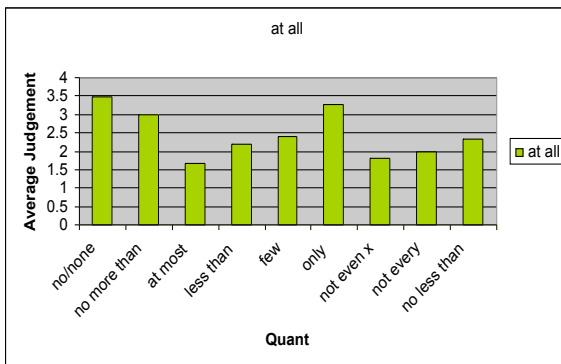
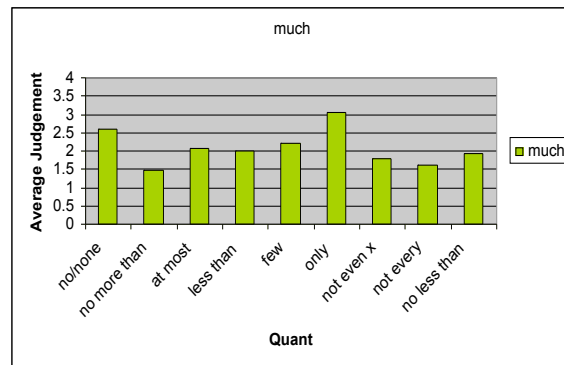
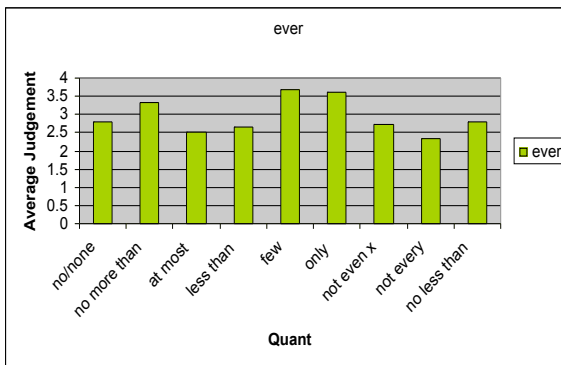
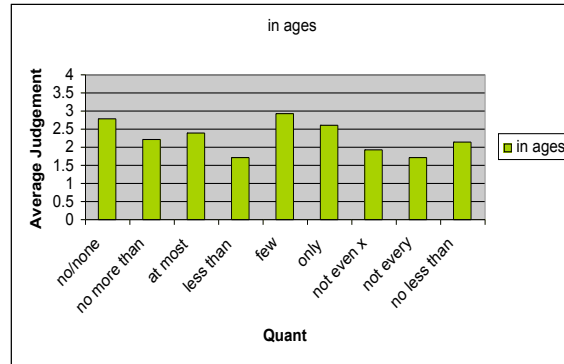
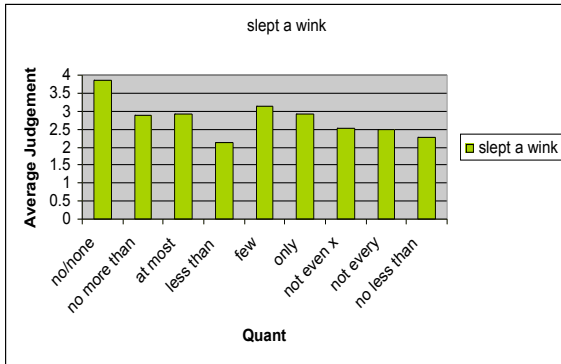
Paul De Decker¹, Erik Larsson¹, Andrea Martin¹

Polarity From Different Perspectives, NYU, March 11—13, 2005

15 participants were asked to rate the grammatical acceptability of 53 sentences (our target sentences are listed in the appendix) using the following scale:

- 1 Completely unacceptable
- 2 Somewhat unacceptable
- 3 Somewhat acceptable
- 4 Completely acceptable

Method #1 NPIs by Quantifiers



1. Dept. of Linguistics, NYU; dedecker@nyu.edu
 2. Eriksen Translations, Inc.; erikl@erikseninc.com
 3. Cognition and Perception Program, Dept. of Psychology, NYU; aem357@nyu.edu

Looking at the graphs for each NPI some patterns emerge:

1. *slept a wink* and *in ages* group together very nicely
2. *ever* should probably also fit into this first group, though it is not as similar to either of *slept a wink* and *in ages* as these two are to one another
3. *much* is somewhat similar to *at all*, but also bears similarities to *slept a wink*
4. *yet* bears similarities to *at all*, but is relatively dissimilar to the remaining NPIs.

This analysis was not based on numbers/actual acceptability averages, but instead on relative hierarchies of quantifiers for each NPI. We arrived at the list above by grouping together the NPIs for which the hierarchies of quantifiers (i.e. the list in descending order from quantifiers most likely to license the NPI to quantifiers least likely to license the NPI) are most similar.

Rather than telling us which NPIs are most acceptable across the board, this analysis tells us which NPIs behaved similarly (i.e. which NPIs are most likely to be licensed by quantifiers a, b, and c, while also being least likely to be licensed by quantifiers x, y, and z), even if the actual acceptability averages are not entirely similar. For example, consider the following quantifier hierarchies for *in ages* and *slept a wink*:

In ages:	Slept a wink:
1. few	1. no
2. no	2. few
3. only n	3. only n
4. at most n	4. at most n
5. no more than n	5. no more than n
6. no less than n	6. not even n
7. not even n	7. not every
8. not every	8. no less than n
9. less than n	9. less than n

The similarities between these lists indicate that these NPIs behaved similarly to one another, i.e. each was far more likely to be licensed by *no* or *few* than by *not every* or *less than*.

The interest of the above findings is twofold.

- (1) The groups of NPIs that emerge are not quite familiar from the standard literature.
- (2) The licensors do not quite group together as the literature would predict. For example,
 - (a) the downward entailing but not antiadditive quantifiers *few*, *less than n*, *at most n*, and *no more than n* do not behave identically;
 - (b) *no less than n*, an upward entailing quantifier, is not an entirely unsuccessful licensor;
 - (c) While *not every* is generally a poor licensor, which can be explained as an intervention effect, it is the No.2 licensor of *yet*. (This is corroborated by many naturally occurring examples found on Google.)

Method #2.

In order to see if there are true differences in the way certain NPIs pattern under these licensors, we ran paired t-tests over the average acceptability scores for each NPI under all licensors. The paired *t*-test focuses on the difference between mean acceptability score (the paired data sets) and reports the probability that the actual mean difference is zero. If the probability (p-value) is less than 0.05, we can reject the null hypothesis that these data sets are the same. Therefore, any $p < 0.05$ is deemed a significant difference, while anything higher is not significant. This kind of t-test assumes the two samples to be compared are not randomly selected: the second data set is the same as the first after some treatment, i.e. the particular NPI, has been applied. If there is a significant difference between the mean acceptability scores between two NPIs, there is reason to believe that they pattern differently from one another; similarly, if there is not a significant difference in two NPIs mean acceptability, then there is no reason to suggest that they should not be grouped together.

Results from our paired t-tests found *no significant differences* in the ratings for:

Slept a wink vs. Ever	p = 0.47	Not Significant
Slept a wink vs. At All	p = 0.09	Not Significant
Slept a wink vs. Yet	p = 0.26	Not Significant
In Ages vs. Much	p = 0.19	Not Significant
In Ages vs. At All	p = 0.33	Not Significant
In Ages vs. Yet	p = 0.33	Not Significant
Ever vs. Yet	p = 0.20	Not Significant
Much vs. At All	p = 0.07	Not Significant
Much vs. Yet	p = 0.13	Not Significant
At All vs. Yet	p = 0.73	Not Significant

Significant Differences: (*indicates $p < 0.05$)

Slept a wink vs. In ages	p = 0.001*
Slept a wink vs. Much	p = 0.003*
In Ages vs. Ever	p = 0.001*
Ever vs. Much	p = 0.001*
Ever vs. At All	p = 0.03*

The mean difference in acceptability score between NPI groups suggest the following two groups:

<i>At all</i>	<i>Slept a wink</i>
<i>Yet</i>	<i>Ever</i>
<i>Much</i>	
<i>In ages</i>	

However, it is important to note that these tests do not investigate the effect of specific licensors on the NPIs or on the average acceptability score.

Appendix: Sentences evaluated

At most two babies have slept a wink since then.
 At most two dogs have barked in ages.
 At most two of my dogs have barked yet.
 At most two of my friends have missed me much.
 At most two people have ever been here.
 At most two people have seen this movie at all.

Few babies have slept a wink since then.
 Few tourists have been here in ages.
 Few roosters have crowed yet.
 Few roosters have crowed much.
 Few students have ever completed the test in less than one hour.
 Few people have read this book at all.

Less than three babies have slept a wink since then.
 Less than three people have seen this movie in ages.
 Less than three birds have flown yet.
 Less than three babies have cried much.
 Less than twenty runners have ever run a mile in under four minutes.
 Less than three boys have hit the ball at all.

No less than five babies have slept a wink since then.
 No less than five professors have given lectures in ages.
 No less than five babies have cried yet.
 No less than five cats have meowed much.
 No less than five people have ever been here.
 No less than five cats have meowed at all.

No more than five people have slept a wink since then.
 No more than five girls have called me in ages.
 No more than two of my professors have given lectures yet.
 No more than five boys have called me much.
 No more than five boys have ever used this bicycle.
 No more than five people have slept at all.

No one has slept a wink since then.
 No rooster has crowed in ages.
 No one has visited me yet.
 None of my dogs have barked much.
 No one has ever seen my invisible friend.
 None of my friends have visited me at all.

Not even five people have slept a wink since then.
 Not even five boys have played chess in ages.
 Not even five people have seen this movie yet.
 Not even five girls have played chess much.
 Not even five doctors have ever performed that operation.
 Not even five professors have given lectures at all.

Not every baby has slept a wink since then.
 Not every cat has meowed in ages.
 Not every car has been sold yet.
 Not every baby has cried much.
 Not every cat has ever meowed.

Not every rooster has crowed at all.

Only five babies have slept a wink since then.

Only two of my friends have called in ages.

Only two of my cats have meowed yet.

Only three people have eaten much.

Only five astronauts have ever walked on the moon.

Only two of my dogs have barked at all.