Not what you expect: The relationship between violation of expectation and negation

Victor Gomes (vgomes@upenn.edu)
Department of Psychology, University of Pennsylvania
425 S. University Ave, Philadelphia, PA 19139 USA

Yubin Huh (huhyubin@sas.upenn.edu)
Department of Psychology, University of Pennsylvania
425 S. University Ave, Philadelphia, PA 19139 USA

John Trueswell (trueswel@psych.upenn.edu)
Department of Psychology, University of Pennsylvania
425 S. University Ave, Philadelphia, PA 19139 USA

Abstract

Language acquisition research has shown that children are delayed in their production and comprehension of truth-functional negation (e.g., "A raven is not a writing desk.") as compared to other kinds of negation (e.g., rejection and nonexistence). The source of this delay is unclear, it may reflect difficulty in mapping the concept of negation to the way it manifests in their language, or it may be due to a lack of a conceptual or cognitive ability. This work aims to investigate the circumstances under which a learner might infer the presence of negation in a message, inspired by the approach of Papafragou, Cassidy, and Gleitman (2007). Namely, we investigate the degree to which videos in which agents fail in completing an action encourages adult participants to infer the use of negation in an utterance describing it. In addition to Event Type (i.e., Failures vs. Successes), we provided participants with additional linguistic information (i.e., syntactic information via Jabberwocky sentences), lexical information (i.e., an alphabetical list of the content words), and Full Linguistic Context (the English sentence with a single item missing). With adults, we ask whether learners with the ability to attend to goals and perceive deviations from their completion could make use of this information, and if so, to what extent do varying degrees of converging linguistic evidence further assist in inferring the use of a negator.

Keywords: language acquisition; cognitive development; numerical cognition; preregistered

Introduction

Early work in psycholinguistics has shown a delay in children's production of truth-functional negation (e.g., "A raven is not a writing desk.") as compared to other kinds of negation such as nonexistence or denial (Bloom, 1968). Recent work has provided support for a delay in comprehension as well, with children younger than 2 years performing at chance on using negative information (e.g., "It's not in the bucket.") (Austin, Theakston, Lieven, & Tomasello, 2019; Feiman, Mody, Sanborn, & Carey, 2017). Since truth-functional negation presents a host of obstacles to a new language learner, ranging from abstractness (e.g., how can one point to or show negation in the world) to salience (e.g., referring to things that are not present or nonexistent), gaining insight into the cues that are available for language learners to exploit for acquiring knowledge of negators in their language is crucial.

Prior research on the acquisition of words with abstract meaning (especially credal verbs, e.g., think, know) has found that linguistic knowledge gives individuals access to helpful evidence for identifying abstract referents, particularly syntax (Gillette, Gleitman, Gleitman, & Lederer, 1999; Snedeker & Gleitman, 2004). Moreover, it is hypothesized that such syntactic evidence would be especially helpful in supportive referential contexts whose conceptual organization align with syntactic structure. Work in this vein by Papafragou et al. (2007) investigated the degree to which particular situations provide helpful cues to reference. They found that specific referential contexts can, on their own, promote the hypothesis that a message might contain a credal verb. In particular, contexts involving false belief (e.g., a woman incorrectly reaching for, and picking up a teapot instead of her teacup) are more likely to be described by participants using credal verbs than the matched, true-belief scenes (e.g., a woman correctly reaching for and picking up her teacup). Interestingly, when participants received converging syntactic evidence (e.g., Jabberwocky sentences like "The zeb ___ that the dax is a florbit.") alongside the false-belief scene, their credal verb responses increased significantly as compared to those who received just scene or syntactic evidence alone. This suggests that a learner armed with linguistic knowledge and the corresponding referent world would be especially good at discovering the meanings of credal verbs in supportive contexts. In addition, the findings suggest that an observer who is capable of inferring the beliefs and/or goals of others would be able to sometimes infer words whose meanings on the surface seem uninferable from the situational context alone (i.e., situations of false belief).

We seek to extend this approach as it concerns negation: Can a learner who is capable of inferring the goals of an agent, develop expectations, and recognize deviations between goals and the resulting state of affairs, infer the use of a negator. Since negation seems similarly unobservable, and, in addition, lacks a specific syntactic form in English unlike credal verbs, an investigation of the relative contributions of linguistic and non-linguistic cues in the inference of negation would provide a better understanding of 1) what kinds of situations tend to elicit negative descriptions and 2) what kinds of linguistic cues are helpful in inferring the polarity of

a statement if one does not have knowledge of a language's negators.

Following previous work that suggested a relationship between fulfillment of expectation and negation (Pea, 1980; Tomasello & Farrar, 1986), we predict that "failure" events (e.g., not catching a ball, not being able to turn on a light) encourage descriptions which include negators. But, some failures can be described affirmatively (e.g., missing a ball) and in other cases there may be other aspects of the scene that can be described affirmatively (e.g., trying to turn on a light), so linguistic cues which draw attention to the failure should further encourage the use of negators.

We report three experiments that investigated the relative contributions of non-linguistic evidence (i.e., visually presented event information) and linguistic evidence (i.e., syntactic and/or lexical information) to inferring the presence of a negator in a message. In all three experiments, each trial consisted of the participant watching a video of an individual carrying out an action, after which the participant was to guess what a parent might say to a child to describe the video. These descriptions were then later coded for the presence / absence of negation. The three experiments always compared two types of events (Event Type): Success events depicted the individual achieving their goal (e..g., catching a ball) whereas Failure events depicted that same individual failing at this goal (e.g., failing to catch the ball). Success versions of an event might involve someone lighting a match, catching a ball, or cutting a leaf, while the Failure versions of those events would show failed attempts (striking, but not lighting a match; missing a ball, or trying and failing to cut a leaf). As mentioned above, Failure events should make negation more salient.

Additionally, in all three experiments, Event Type was crossed with a second factor (Information Type), such that participants provided their response to Success and Failure events under one of two conditions pertaining to the kind of information provided about the parent's message. In the Video Only condition (which was present and the same in all three experiments), participants only saw the video and received no linguistic clues about what the parent had said, and thus had to provide their own sentence (with or without negation) to describe the video. In the Video+Language condition, participants were provided with additional clues about what the parent had said. The kind of linguistic information differed across experiments. In Experiment 1, participants in the Video+Language condition were provided with syntactic clues about the message; a Jabberwocky version of the "parent's utterance" was presented below the video (e.g., "The rizz did ___ bleck the dax," for "The girl did not catch the ball."). Participants were asked to provide the missing word in English. In Experiment 2, those in the Video+Language condition were instead provided with a list of content words that were present in the target utterance (e.g., below the ball-catching video, an observer was provided with "ball, catch, girl", which is an alphabetically ordered list of the content words found in "The girl did not catch the ball."). Participants in this condition assembled a sentence using these words plus any additional words they thought were needed. Finally, in Experiment 3, those in the Video+Language condition were provided with the full linguistic context, i.e., an English sentence with a missing word that they were told came from parental descriptions of the scenes (e.g., "The girl did ___ catch the ball."). They were asked to provide the missing word in English.

Thus, all three experiments had a 2x2 design consisting of one within-subject factor (Event Type: Success vs. Failure) and one between-subject factor (Information Type: Video vs. Video+Language). If as predicted, Failure events lead learners to infer the presence of negation in a message, all three experiments should show a reliable main effect of Event type, such that Failure events will elicit more negation than Success events. If linguistic evidence from the message also helps infer the presence of negation, then one would expect a reliable interaction between Event Type and Information Type, such that the effect of Event Type (Success vs. Failure) should be larger when additional linguistic evidence is provided (Video+Language) than when it is not (Video). Finally, we ran an additional base-line (Language Only) condition for all three experiments. Participants in this condition never saw the videos and instead answered using only the linguistic clues. This permits us to examine the contribution of linguistic evidence on its own in the absence of a scene.

Method

Preregistration

The preregistration can be accessed at: osf.io/wt6p2. One deviation occurred; twice as many people participated than as planned due to a miscommunication between lab members. This deviation was discovered only after the experiment was complete. Analyses below include data from all participants.

Participants

A total of 112 English-speaking participants were recruited on MTurk and completed the task on PCIbex (Zehr & Schwarz, 2018), of which 4 were dropped (1 for providing ungrammatical responses and 3 for not following instructions) and replaced. Each experiment had 36 participants, 12 in each condition (i.e., Video; Video+Language; and Language Only).

Materials

Visual materials 44 short videos (each under 20 seconds) of people performing simple actions were recorded and used in the experiment. Of the 44 videos, there were 20 filler items and 24 experimental items. For the experimental items, 12 actions were used in the 24 items, with each action having two video versions, one in which the action succeeds (e.g., a person catches a ball) and another in which it fails (e.g., a person misses a ball). Experimental items all begin similarly, and only diverge minimally at the point of completion.

Linguistic materials The sentences used in the Language conditions of the experimental items are based on participant descriptions of failure scenes from a pilot version of this experiment run on different participants. For example, one vignette involved a person lighting a match (i.e., the success version) and a similar video in which the match did not light following the person's attempt (i.e., the failure version). In the pilot, participants often described the failure version of the scene as "A person can't light a match." Participants in the Jabberwocky condition would see "A tive can ___ glizz a dax," and those in the Content Words condition would see "light, match, person," And finally, for the Full Linguistic Context condition, they would see "A person can ___ light a match." The nonsense Jabberwocky words were partly derived from previous experiments (Gillette et al., 1999; Snedeker & Gleitman, 2004) as well as the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002). For filler items, the linguistic stimuli were developed to resemble the experimental item stimuli, but other parts of speech were randomly selected to be removed (either adjectives, nouns, adverbs, verbs, prepositions, and conjunctions).

Procedure

For all experimental items with video, Event type was a within-subject condition. In particular, participants were randomly assigned to a list in which half of the events were Success scenes and the other half were Failure scenes. Lists are matched such that participants only saw one kind of event per video (e.g., if one saw the Success version of an event, one would not see the Failure version). For each experiment, participants were randomly assigned to one of the two Information type conditions (Video; Video+Language) or to the baseline (Language-Only) baseline condition. For all three experiments, participants in the Video Only condition saw the video and were told "You are about to see a series of silent videos. We asked parents to describe these videos to their children, and your task is to guess how a typical parent might have described it to their child." Those in the Video+Language condition got different instructions depending on the Experiment. Those in Experiment 1 (Syntax) were told "You are about to see a series of silent videos and sentences that English-speaking parents have used to describe these videos to their children. These sentences have been translated into a novel language. We have removed ONE WORD from the sentence. For each video, we want you to provide the missing word that you think a parent might have used." Those in Experiment 2 (Content Words) were told "You are about to see a series of silent videos and a list of words from sentences that English-speaking parents have used to describe these videos to their children. After each video, you will be presented with a list of words (in alphabetical order) that parents used in describing the video to their children. For each video, we want you to provide a sentence that you think a parent might have used." Those in Experiment 3 (Full Information) were told "You are about to see a series of silent videos and sentences that English-speaking parents have used to describe these videos to their children. We have removed ONE WORD from the sentence. For each video, we want you to provide the missing word that you think a parent might have used."

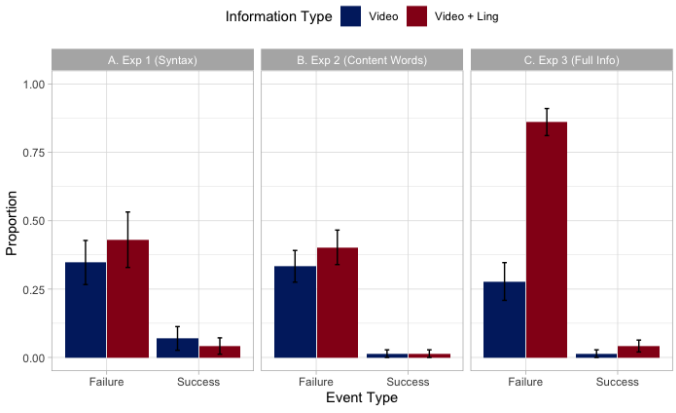

Figure 1: Figure 1: Proportion of negators (*no*, *not*, *-n't*) included in response by Information Type and Event Type for all experiments. Average of Subject means. Error bars indicate ±1 Standard Error.

Results

Experiment 1: Syntax Figure 1A presents the proportion of trials for which a negator was inferred, split by Event Type (Success vs. Failure) and Information type (Video vs. Video+Language). As can be seen in the figure, participants on average were more likely to infer negation from situations of Failure (38.9%) than Success (5.6%) collapsed across Information Type. Moreover, the addition of Syntax (in the Video+Language condition) did not seem to affect the ability to infer negation from Failure vs. Success events. These conclusions were supported by a multilevel logistic regression that predicted negation, using Event and Information type, and their interaction as fixed effects. (In all models reported, a maximal random effects structure for subjects and items were used, simplifying for nonconverging models). We found a reliable main effect of Event type (β=4.011, SE=1.501, z=2.672, p=0.008), no effect of Information type (β=-0.149, SE=0.915, z=0.162, p=0.871) and no reliable interaction (β=1.379, SE=1.487, z=0.928, p=0.353). The four conditions plotted in 1A were also individually compared to the base rate (8.3%) of inferring negation from the Jabberwocky sentence alone (Language Only condition). Only the Failure condition of the Video Only condition differed from base rate (β=3.679, SE=1.471, z=2.501, p=0.012), though there was a trend when comparing the Failure condition of the Video+Language condition to the base rate (β=10.93, SE=5.825, z=1.876, p=0.061). When participants failed to produce negations in the Jabberwocky + Video and Jabberwocky alone conditions, they most frequently (18%) produced *-ly* adverbs (e.g., *quickly*, *completely*, *easily*).

Experiment 2: Content Words Figure 1B presents the proportion of trials for which a negator was inferred in Experiment 2, split by Event Type and Information type. As can be seen in the figure, participants on average were more likely to infer negation from situations of Failure (36.8%) than Success (1.39%); the addition of the language (Content Words) did not seem to affect the ability to infer negation. These conclusions were supported by a multilevel logistic regression[1]. In particular, we found a reliable main effect of Event type (β=4.014, SE=0.786, z=5.107, p<.001), no effect of Information type (β=-0.1927, SE=0.8304, z=-0.232, p=0.817) and no reliable interaction (β=0.7739, SE=1.5514, z=0.499, p=0.618). The base rate of guessing negation from just the Content Words alone (the Language Only condition) was 7.6%. Only the two Failure conditions were different from this base rate (Video: β=1.705, SE=0.494, z=3.449, p<0.001; Video-+Language: β=2.082, SE=0.679, z=3.063, p=0.002).

Experiment 3: Full Linguistic Context Figure 1C presents the proportion of trials for which a negator was inferred in Experiment 3, split by Event Type (Success vs. Failure) and Information type (Video vs. Video+Language). As can be seen in the figure, participants on average were more likely to infer negation from situations of Failure (56.9%) than Success (2.8%). Here, the addition of the full linguistic context seemed to affect the ability to infer negation from these scenes. These conclusions were supported by a multilevel logistic regression that predicted negation, using Event and Information type, and their interaction as fixed effects. In particular, we found a reliable main effect of Event type (β=5.520, SE=0.985, z=5.605, p<.001) and Information type (β=2.228, SE=0.918, z=2.426, p=0.015) plus a reliable interaction (β=3.735, SE=1.825, z=2.046, p=0.041). Here the base rate of inferring negation from the linguistic context (Language Only) condition was much higher than the previous experiments 45.6%. The two Failure conditions differed from this base rate (Video: β=-1.208, SE=0.002, z=-501.364, p<.001) (Video+Language: β= 2.908, SE=0.968, z=3.004, p=0.003). There was no reliable difference from base rate for the two Success conditions, with the model failing to converge in the Video+Language condition, and trending for the Video Only condition (β=-6.416, SE=3.675, z=-1.746, p=0.081). When participants failed to produce negations in the Full + Video and Full alone conditions, they typically (70%) produced *-ly* adverbs.

Pairwise comparisons between Experiments Models comparing the Failure scenes of all three experiments to each other (i.e., 1 to 2, 2 to 3, and 1 to 3) found no reliable effects or interactions when comparing Jabberwocky to Content Words, but when comparing either of the two to the Full Linguistic Context.

The model comparing Experiments 1 (Jabberwocky) and 2 (Content Words) found no significant effects of experiment (β=-0.04108, SE=0.38263, z=-0.107, p=0.914), Information Type (β=0.61744, SE=1.22175, z=0.505, p=0.613), and no reliable interaction (β=-0.14199, SE=0.76551, z=-0.185, p=0.853).

Comparing Experiments 1 (Jabberwocky) and 3 (Full Linguistic Context) found significant effects of experiment (β=1.33949, SE=0.54261, z=2.469, p=0.0136), Information Type (β=2.23286, SE=0.56131, z=3.978, p<0.001), as well as a reliable interaction of both (β=3.29817, SE=1.10333, z=2.989, p=0.003).

Finally, Comparing Experiments 2 (Content Words) and 3 (Full Linguistic Context) found significant effects of experiment (β=1.14023, SE=0.35250, z=3.235, p<0.001), Information Type (β=1.77007, SE=0.36678, z=4.826, p<0.001), as well as a reliable interaction (β=2.84034, SE=0.71805, z=3.956, p<0.001).

Discussion

As we initially predicted, Failure events were significantly more likely to promote negation responses than Success events across all Experiments. This effect of Event Type was enhanced by the addition of the full linguistic context (Exp 3), but not by the addition of just syntax (Exp 1), or content words (Exp 2). The presence of a main effect of Event Type in all three experiments suggests that there may indeed be particular situational contexts in which a learner could infer the presence of negation in a message, even when that learner does not have access to any linguistic information about that message (as evidenced by the effect of Event type within the Video Only conditions). In a context in which a speaker is commenting on a situation in which someone failed to achieve a goal, a learner might reasonably posit that this utterance contained negation. Note that this requires a learner who is sophisticated enough to interpret the goals of others and evaluate when those goals were not satisfied, as well as something about the pragmatics of language use. In this regard, it is notable that work within social development suggests that even 18 month old humans attempt to help adults who are unable to accomplish a goal (Warneken & Tomasello, 2006, 2009). Thus, it is possible that young children might be able to use failure contexts (like the ones used here) to aid in the learning of negation, although the literature to date suggests delays in this ability.

Interestingly, a slightly more sophisticated learner who has access to a syntactic evaluation of the utterance (as simulated in Exp 1) or to the meanings of the other words in the sentence (as simulated in Exp 2) appeared to obtain no additional benefit for inferring negation, above and beyond what can be gleaned from the situational context alone. (In particular, there was no reliable interaction between Scene Type and Information type in either Exp 1 or Exp 2). We predicted

[1] In Experiment 2, the Success Video of the Video Only condition had no negation responses at all (0%), which caused the generalized linear model (glmer in R) to fail to execute. We decided to randomly select a single trial from a single subject and change the observed value from 0 to 1, which permitted the model to converge without changing any of the observed patterns reported here (the mean went from 0% to 1% as shown in the figure)

this to be the case for syntactic information (Exp 1) since at least in English there are no strong syntactic cues to negation being present in a sentence (beyond the presence of the negator itself). However, we hypothesized that knowing the content words in an utterance might help a speaker identify which event construal was being labeled (e.g., door, woman, open) but it appears that the addition of this information was insufficient. We now think that this may be due to the fact that a participant/learner could have used these same content words to label some other event in the scene (e..g, the fact that some other woman was able to open the door before the girl who failed).

The fact that partial access to the linguistic context (Exp 1 and Exp 2) did not enhance the detection of negation provides us with an additional hypothesis for why children show delays in understanding negation. Young language learners (who likely only have partial access to the parse of the sentence and/or the meanings of the other words) would struggle to identify negation from these situations of use. Indeed, it was only a truly sophisticated language learner/user (as simulated in Exp 3) who shows a significant advantage in inferring the presence of negation. Here, probabilistic evidence derived from the utterance itself supports negation (e.g., "The girl could ___ open the door."), resulted in 45.6% negation responding without any scene. If we accompany such a sentence with a situational context of Success or Failure, now it is relatively easy to infer when negation is, or is not, intended by the speaker (as evidenced by 5% negation responses for Success events vs. 80% for Failure events). The effect of linguistic information requiring the combination of syntactic and lexical information is consistent with the fact that negation interfaces with syntax (i.e., scope) and does not occupy a unique syntactic position in English. In other words, the lexical information helps narrow down possible referent events in the world, and the syntactic information provides a frame which constrains candidate sentence forms (e.g., excluding the use of "trying to").

In sum, we found initial support for our prediction that learners who are able to infer goals, develop expectations, and notice deviations from them, would be able to use this information to infer the presence of a negator in a description of that scene. Additionally, neither syntactic information alone nor lexical information alone has an effect. This has potential consequences for work in children's acquisition of negation, as children, if they are equipped to infer goals, develop expectations, and notice deviations from them, may additionally be able to use this information to assist them in learning the negators of their language. It additionally may speak to the results mentioned above concerning the observed delay in children's comprehension of negation before age 2, as children typically exit the two word stage of production after age 2 (the stage at which most utterances are comprised of two words, e.g., "Mommy eat."), which strongly indicates a meaningful growth in syntactic abilities (Bloom, 1968). If this is the case, then the delay may be because they are mostly only able to make use of lexical knowledge and are limited in the ability to fully exploit the syntactic information that is also available to them in supportive learning contexts.

Additional work is underway to address limitations of the present study and broaden our conclusions. Most notably, the referential contexts that support negation were artificially generated based on intuition, and there may be other contexts (e.g., contrast, or absence) that similarly encourage the inference of a negator in a message. If our conclusions have any merit, it is urgent to evaluate the situations of use of negation in natural environments, where parents spontaneously produce negation in the home. Our work to date indicates that negation is common in parent's speech to children. We are now examining the extent to which the linguistic and nonlinguistic contexts of use show effects of informativity similar to those found here.

References

Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2019). Young children's understanding of denial. *Developmental Psychology*, *50*(8), 2061.

Bloom, L. M. (1968). *Language Development: Form and Function in Emerging Grammars*.

Feiman, R., Mody, S., Sanborn, S., & Carey, S. (2017, October). What Do You Mean, No? Toddlers' Comprehension of Logical "No" and "Not". *Language Learning and Development*, *13*(4), 430–450.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999, December). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176.

Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition*, *105*(1), 125–165.

Pea, R. (1980). *The development of negation in early child language* In Bruner S. & Olson D. (Eds.), The Social Foundations of Language and Thought: Essays in Honor Of Jerome. New York: WW Norton.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The arc nonword database. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1339–1362.

Snedeker, J., & Gleitman, L. (2004). Why it is hard to label our concepts. *Weaving a lexicon*, *257294*.

Tomasello, M., & Farrar, M. J. (1986). Object permanence and relational words: A lexical training study. *Journal of Child Language*, *13*(3), 495–505.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *science*, *311*(5765), 1301–1303.

Warneken, F., & Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends in cognitive sciences*, *13*(9), 397–402.

Zehr, J., & Schwarz, F. (2018). *Penncontroller for internet based experiments (ibex).* https://doi.org/10.17605/OSF.IO/MD832.