

A one-stage distributed processing account of linguistic negation

Stephanie Huetten (shuetten@ucmerced.edu),

Sarah Anderson (sarah.anderson@ucmail.uc.edu),

Department of Psychology, University of Cincinnati, Cincinnati, OH, 45221, USA

Teenie Matlock (tmatlock@ucmerced.edu),

and Michael J. Spivey (spivey@ucmerced.edu)

School of Social Sciences, Humanities & Arts, University of California, Merced, Merced, CA, 95344, USA

Abstract

Linguistic negation can be comprehended with the inclusion (or absence) of features and categories associated with the senses in a single step. Under this view, there is no need for explicit logical operators, as the negating word or phrase is treated no differently than any other word. Negation provides additional context, whereby visualizing negation as a trajectory in a distributed, grounded perceptual simulation space can easily characterize the comprehension of negated sentences. A mousetracking experiment was conducted to explore how this kind of process may be enacted in the brain and to tease apart hypotheses of logical manipulations vs. analogue signals performing this work.

Keywords: language understanding, neural networks, situated cognition, mouse-tracking

Introduction

Negation occurs in all languages and is an important part of everyday communication. Humans need it to express situations in which an object, person, state, or event is not present or even non-existent. Negation can be coded in a variety of ways. In some cases, it is expressed suprasegmentally, for instance, in sarcastic remarks such as “I sure excelled in that class,” said by a student who just learned that his course grade was a D. It can also be coded with affixes, such as the prefix “non-“, as in “The nonexistence of rules worries me.” It can also be coded with words such as “not”, “none”, and “without”. Despite the ubiquity of negation, its processing has perplexed philosophers for centuries. There is still a relatively poor understanding of how people contend with the meaning of simple negated statements such as “The eagle is not in the sky”.

The processes that underlie the comprehension of a negated utterance has generated recent interest among psycholinguists. In pioneering work, Kaup, Yaxley, Madden, Zwaan, and Lüdtke, (2006) presented participants with sentences like, “The eagle was not in the sky,” then presented an accompanying picture. Participants had to respond whether the content of the picture had been mentioned in the sentence they had just read. The reasoning was that if participants perceptually simulate negation in the same way they simulate affirmative sentences, they should be quicker to respond “yes” to pictures that match the sentence (e.g., an eagle with its wings folded). Kaup and colleagues found that when participants read negated sentences, response times to a picture that matched the

affirmative version of the proposition (eagle with wings spread) were faster than to a picture that actually matched the negated sentence, suggesting that a perceptual simulation of the affirmative proposition was created in response to negated sentences. The researchers then suggested that negation may be handled through a separate, special process, not through a perceptual simulation.

Thus, the timecourse of understanding a negated sentence is not well understood. There are three possibilities as to how a final interpretation of the negation comes to be activated (Figure 1). The first possibility is a logical process discretely revises a fully activated affirmative interpretation (1A). The second scenario predicts a lag in activation of a negated sentence, simply because negation is more complex (1B). The third hypothesis is negation produces increased competition amongst the possible features the negated meaning could activate, forcing negation to be slower at first, but once one interpretation begins to form, making negation much faster compared to the affirmative version (1C). Note that this last hypothesis predicts only speed differences, and no spatial differences.

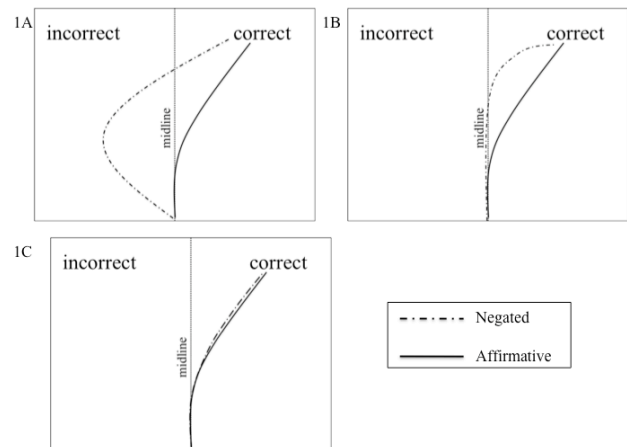


Figure 1: Each hypothesis shows predicted mouse trajectories on a computer screen.

Each of these panels in Figure 1 can be thought of in terms of velocity: the speed and direction in which an interpretation is formed will be shared with motor systems during a response (Spivey, Grosjean & Knoblich, 2005). Thus, spatial and velocity metrics of a computer mouse will reveal which of the aforementioned hypotheses can best explain how a negated interpretation is formed. In addition, a proof of concept model was created as a fundamental first

step in providing a groundwork for a one-stage process for linguistic negation. Although preliminary, proof of learning and some principles of perceptual simulation are demonstrated.

Inspired by Elman's (1990) treatment of simple recurrent networks that learn linguistic sequences, one can conceptualize the dynamics of a negated sentence as forming a trajectory through the state space of the network. If each neuron is treated as a dimension, then the firing rates of all the neurons at any given time slice form a set of coordinates that index a location in the state space. Therefore, a string of words in a sentence necessarily forms a string of locations: a trajectory. Importantly, sensorimotor features can be integrated into a simple recurrent network (Howell, Jankowicz, & Becker 2005), creating dimensions that arise from perception and emerge in cognition. Simply put, the meaning of a word in this network is a distributed code of features and primitive perceptual categories.

Whereas some previous networks use a single node with a priori instructions on performing the negation (e.g., Samad, 1988), this network has no special instructions for the word *not*. Instead, when the network is presented with the words *The coin is not heads-up*, the network sees features on the last word that are instead associated with features on a tails-up coin. Thus, *not* becomes a contextual modifier, signaling that the upcoming word is of an alternative or more diffuse meaning. More specifically binary-alternative negations can readily go to the alternative, and multiple-alternative negations (such as location, in which an object could be in any number of alternative places) become a blend of possibilities.

First the network will be laid out and its properties examined for principles of how perceptual simulation must function under a framework of embodied semantics. Then, a mousetracking experiment will demonstrate that a one-stage process, with increased competition for negated sentences, best accommodates these and previous findings.

Model

The corpus consists of both affirmative and negated sentences, in even numbers. Eleven sentences were binary-alternatives (e.g. The coin is [heads-up/tails-up]) and 12 sentences were four-alternative (e.g. The child is [angry/annoyed/calm/content]). Some partial overlap was built into this corpus (i.e., with two of the binary-alternative pairs using glass as their noun). It is much less clear how the network would negate a multiple-alternative set in this case, and thus their examination is crucial to test the flexibility of this model. Further, some overlap must realistically be built in to mimic a realistic scenario. Learning rate was kept constant at a value of .09 and momentum was set at a value of .1 for the duration of training.

The target perceptual features were constructed using binary targets, where the first-author rated each word as either having a particular feature or not. The features chosen were visual (33 features), proprioceptive (11 features), auditory (5 features), emotional (8 features) or related to

olfaction/gustation (6 features). These 63 features targets were specified a priori but do not imply that these features are exactly their label, or must be in some predetermined arrangement. The target layer can be thought of as the environment and the things that are simply co-present with a particular word, and the important aspect is on the feature-feature relationships, and not the label itself. The network was trained such that error consistently ranged from the minimum error of 1.9×10^{-7} to a maximum of .026. This was found to be the point at which the error became stable and could not improve any further.

No error was fed back from the perceptual simulation layer on *the*, *is* and *not*. The network is doing word prediction, but it is also seeing and hearing features for the current word. Thus, the target layer for the input *eagle* in the sentence *The eagle is flying*, the corpus target predicts *is* and the perceptual simulation targets are, for example, feathers, eyes, and blue. Thus, the model is not doing any explicit prediction in the perceptual simulation layer. Rather, the most sophisticated assumption is that one needs to have some feature detectors and a modicum of perception to learn about words.

Each axis is a particular feature's activation, chosen automatically. The x-axis is always the most active noun feature for that sentence. The y and z-axes are chosen based on which features are non-overlapping perceptual targets. However, note that overlap between noun and adjective features do not always clearly delimit these boundaries. For example, Figure 2 is an illustrates the way the binary input string *The shoe is (not) [tied/untied]* traverses through a three dimensional coordinate feature space (stable, tough, moving). *The shoe* follows the same path for every sentence, going from stable, to stable and tough, to a location with partially active stable and moving features. This is already displaying an emergent property: the model was not told to activate any features for *the*, *is* and *not*; however, these words still carry some feature activations. Also note that the context layer of the network is reinitialized after each sentence is run through the network, and so activation for *the* cannot be due to carry over from the previous sentence. Due to the pattern of connections, these words cannot help but carry multifarious meaning.

After *is* in Figure 2, the trajectories begin to diverge. Notice that both negated sentences overlap on *not* and that this location is approximately equidistant between two alternative locations. *The shoe is tied* and *The shoe is not untied* both end in a corner with stable being almost entirely active. *The shoe is untied* and *The shoe is not tied* end by deactivating stable and tough features, and activating moving features. Thus, when the network hears about a shoe being untied in either its negated or affirmative versions, it is able to comprehend a stable tough shoe, and then correctly simulate something akin to shoelaces moving around freely as one walks.

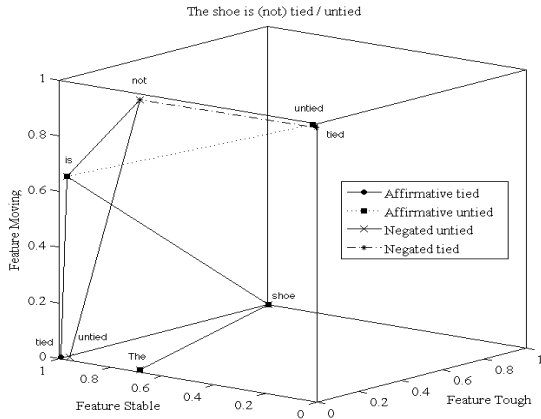


Figure 2: Sentence trajectory through feature activation space. Beginnings of sentence overlap until the noun.

Figure 3 shows some nuance of how negating a sentence may subtly change its meaning. *The shoe is tied* settles easily into a perceptual attractor, whereas *The shoe is not untied* shows a slightly different pattern of activation. The targets did not expressly tell the network that these two meanings were different in any way. However, looking at these trajectories, it appears that the network emergently decided that this particular negation should not activate the *happy* feature that the equivalent affirmative meaning activates. Thus, this network is capable of negating sentences and finding a meaning where primary features are extremely similar in their activations, but other less active features are susceptible to quirks, resulting in subtle, but quantifiable, flavoring differences. The tiny activations surrounding the major primarily activated features create the majority of flexibility in this system, while the primary features constrain and ensure similar interpretations for each sentence.

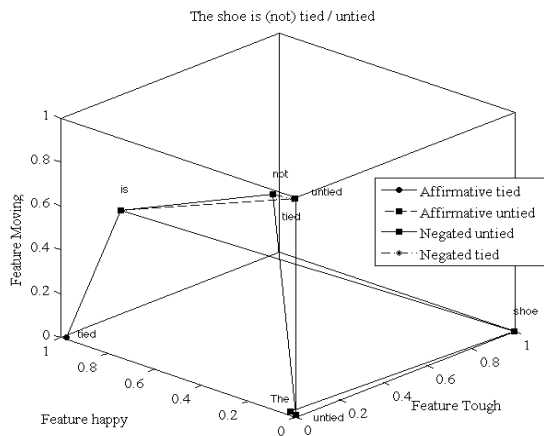


Figure 3: Same sentence as Figure 2 with modified y-axis.

It is much more difficult to see the primary activations of multiple-alternative negations in just three dimensions.

Figure 4 shows a pair from the negation of the trained multiple-alternative sentence (*The computer is (not) [blue, green, gray, white]*). The network has learned to associate *green* primarily with a tangy taste, and *blue* with quietude. However, the negations go to 0, because their alternative meanings are not within this subset of features but would require plotting alternative feature activations. For this visualization, their meaning is in the absence of the primary distinctive features of *green* and *blue*. One of the differences between multiples and binaries lies in this diffusion of meaning. It will always be more difficult to capture the meaning of a negation of something like a location, where the range of alternatives is large or even infinite, or something like a color, which has finite labels but is a continuous physical substrate. Therefore, the meaning of *not* may actually primarily be a deactivation of the affirmative version's primary features and subtle activation of diffuse alternatives. This makes any multiple-alternative negation much less stable, which could affect the encoding and subsequent retrieval accuracy of a negation (Anderson, Huette, Matlock & Spivey, 2010).

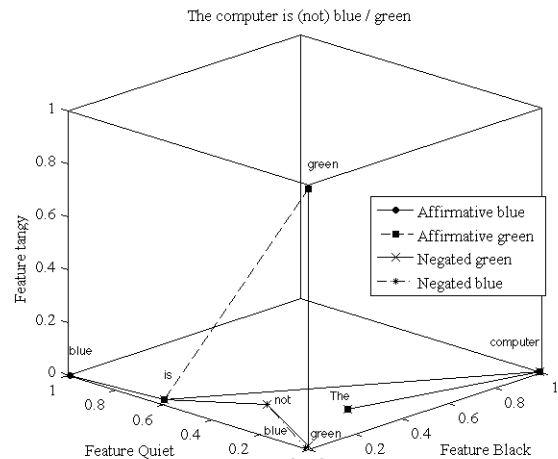


Figure 4: Deactivation in negated meaning

Even in this extremely small corpus, the network is capable of doing some degree of generalization, which previously has been a criticism of the perceptual simulation framework. For example, the network has never seen the sentence *The egg is clean* or *The egg is dirty*. It has encountered each one of those words though, and has associated features. Figure 5 shows appropriate perceptual features for these two novel affirmative sentences, but some less appropriate feature activations for the negated versions of these novel sentences. Though imperfect, the endpoint of *...is not dirty* is closer to the endpoint of *...is clean*, and the endpoint of *...is not clean* is close to the endpoint of *...is dirty*. As such, the previous graphs of oft-repeated sentences may be somewhat idealized versions of how a sentence is negated.

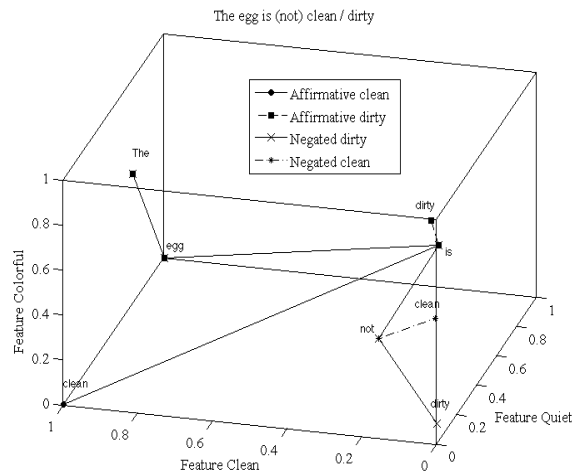


Figure 5: Example generalization in the network

Discussion

To allow for generalization, the meanings of sentences may be necessarily somewhat distinct between the affirmative and the (supposedly) equivalent negated form. Generalization is extremely powerful and indicates that once a small amount of statistical information is encountered and learned, almost anything can be negated as effortlessly as something that has been episodically encountered. When the network encounters a novel use of a word, it can easily simulate it, incorporating the surrounding context to create a brand new meaning. The productivity of language is a fundamental law in linguistics (REF?), and this model and framework holds promise for allowing a multitude of meanings that have never been directly encountered *as well as their negations*.

Experiment

Participants. Twenty-four participants were given partial course credit for their participation and run in accordance with IRB regulations. All were University of California, Merced, students, and all were prescreened to ensure normal or corrected to normal vision and audition. Participants with dyslexia or other reading disabilities were also excluded. Only right-handed or ambidextrous native American English-speakers were included in this study.

Materials. All target sentences were recorded in the negated form, and all fillers were recorded exclusively in the affirmative form. In both cases, a male speaker was used. To construct the affirmative versions, *not* was spliced out of the file at zero crossings, ensuring the voice sounded unaltered and natural. This was done to ensure there were no effects of priming, due to sentences being presented in both negated and affirmative forms. All pictures were simple clipart, modified in Adobe Photoshop to fit the location and state in which they were described. For example, a folded newspaper was presented on a rack, paired with the

sentences *The newspaper is [not] on the rack* or *The newspaper is [not] folded*.

Design. To test whether the affirmative state is competing with or being activated in parallel with the negated state, participants completed a modified forced-choice mousetracking task. Twenty-four stimuli pairs were constructed, where 12 pairs described locations of a noun (*The towel is [not] on the bar/floor*) and 12 pairs described states (*The towel is [not] flat/crumpled*). Two lists were constructed, whereby 6 pairs of each type were used as targets in List A, and the other 6 pairs of each type were used in List B (12 pairs, or 24 individual stimulus sentences). The targets consist of both Negated and Affirmative forms, 2 different states/locations, yielding 48 targets. In addition, filler sentences with a different noun and location or state were constructed, such that the filler trials could consist of two choices with no overlapping information. The pairs not used as targets were used to construct these, randomly selecting from one of the state/location sentence endings. For example, a target trial contained a picture of a newspaper on a rack paired with a newspaper on a driveway, and the comparable filler trial paired a sentence describing a newspaper as folded (same picture of it folded on a rack) with a football on a field. Thus in list A, the newspaper was always presented as a pair of pictures for affirmative and negated spoken stimuli about the location, and the newspaper, described as folded, was always presented with audio and visual components of *The motorcycle is broken* (also a description of its state). This kind of pairing was chosen such that the key manipulation of affirmative and negated was less salient, because sometimes the disambiguating information is in the noun (e.g. picture of a football and a newspaper), and at other times participants had to disambiguate based on the state or location (e.g. newspaper on driveway or rack).

Procedure. The experiment began with five practice trials. Participants read instructions that told them to click a small red box located at the bottom middle of the screen. After clicking this box, it disappeared and a sentence played over headphones. After the sentence, two pictures appeared in the upper right and upper left parts of the screen. The location of the target picture was pseudo-randomized for each list. Participants were asked to move their mouse over one of the pictures to indicate which picture best matched the sentence (e.g. “The newspaper is on the rack” and a picture of a newspaper in an open newsrack or with a picture of a newspaper on a driveway as the response choices). Moving the cursor to the top right or top left over the picture constituted a response (no click was required). X,Y screen coordinates were sampled at approximately 67 Hz using Psyscope X (Cohen, MacWhinney, Flatt & Provost, 1993), along with the items on the screen, reaction times, and the eventual response. The tracking speed of the mouse was set to the third lowest setting in the Mac OSX operating system preferences, and chair height was set to the highest setting. Both of these constraints ensure motor movements come from the entire limb and are not localized to small wrist

movements, which could cause asymmetries in leftward and rightward movements.

Mouse-Tracking Data. Variability in RT led to a different number of samples on each trial. To cut out this variability, each trial was time-normalized in increments of 2% of the total response time by means of linear interpolation (Spivey et al., 2005). This yielded 51 x,y samples per trial, which were then translated to a common starting point of (0,0). All responses to the left side of the screen were mirrored to the right such that all data could be pooled. To derive maximum deviations, a straight line from the start (0,0) to the response (last sample) was computed, and 51 points along that line. Then, Euclidean distance was found from every point on that line to every point on the trajectory (51 points on the line x 51 points on the trajectory equaled 2,601). Because Euclidean distance is an absolute value, points that deviated below the straight line were multiplied by -1, ensuring these points would not be interpreted as deviation toward the competing response. Then, we computed the minimum distance for each point and the maximum of those minimums is found. The resulting point from the trajectory is the maximum deviation.

Results

Importantly, only correct trials were analyzed. Response accuracy was high for all participants, but there was a slight difference in accuracy by condition. For Affirmative trials, an average of 97% were correct ($SD=3.5\%$), and for Negated trials, the average was significantly lower at 93% ($SD=6.5\%$, paired samples t-test: $t(23)=3.08$, $p=.005$). As in previous research, Negated trials exhibited longer reaction times than Affirmative trials, confirmed by a paired samples t-test (Negated: $M=3179\text{ms}$, $SD=671\text{ms}$; Affirmative: $M=2608\text{ms}$, $SD=442\text{ms}$; $t(23)=-7.96$, $p<.0001$). This finding is in line with previous research, aiding in the validation of this as a useful and informative design.

The tests designed to address the three hypotheses were as follows: The proportion of time spent over the midline (Figure 1A), spatial deviations in the form of maximum deviation (Figure 1B), and instantaneous velocity, which should parallel one of the panels in Figure 1. Importantly, the third hypothesis (Figure 1C) only predicts velocity differences as a function of negated sentences competing with an affirmative simulation, but no spatial deviations.

Slight deviations over the midline created a rather high proportion of total time spent on the side opposite the response. The proportion of time spent on the incorrect side of the midline for Negated stimuli was .317 ($SD=.09$) and for Affirmative stimuli the proportion was .29 of the total time ($SD=.1$). This difference is non-significant ($t(23)=-1.307$, $p>.2$). Further, these proportions may seem high, but to see if these were slight deviations or rather large attractions over the midline toward the incorrect response, a threshold of 10 pixels over the midline was used. In this analysis, the proportion of time spent more than 10 pixels past the midline in the wrong direction is .15 for both

conditions (Affirmative $SD=.12$, Negated $SD=.1$, $t(23)=.117$, $p=.91$). This analysis demonstrates that these are slight deviations not to be confused with attraction toward a competing representation.

The next analysis uses the maximum deviation metric to investigate spatial deviations between conditions. This measurement is the maximum pixels deviated from a straight line to the response. (Affirmative trials the mean was 285 ($SD=184$) and Negated mean was 299 ($SD=189$). This difference was non-significant ($t(24)=-.775$, $p=.446$.) This lack of spatial differences in combination with a lack of deviation over the midline, though null results, begin to point toward evidence for a parallel, one-stage process, not a process of revision.

Instantaneous velocity was computed, shown in Figure 6 by condition. This is simply the distance moved between one normalized timestep and the next. Therefore, along the x-axis, the 51 original timestamps become 50 time difference increments. On Negated trials (dashed lines), participants ramped up their speed at a slightly slower rate than on Affirmative trials (thus somewhat slower acceleration for Negated trials around timesteps 42 through 45). Wojnowicz, Ferguson, Dale, and Spivey (2009) previously reported this pattern of slower acceleration for computer-mouse movements in trials that involved competition between simultaneously activated response options. Moreover, when they used Usher and McClelland's (2003) differential equations to model this dynamic competition, they found that the slower acceleration during fierce competition was followed immediately by an abrupt bifurcation of activation profiles. This bifurcation caused an increased spike in activation dynamics in the simulation, and correctly predicted an increased spike in velocity for computer-mouse movements in the high-competition trials. Similar to Wojnowicz et al., the greater competition (and slower acceleration) in Negated trials (Figure 6) is immediately followed by a spike of greater velocity (timesteps 46-49). Essentially, when a negated sentence generates competition between the features of a perceptual simulation, that competition initially impedes the acceleration phase of the motor movement, but once the competition is resolved, the winning response alternative gathers extra speed in approaching its movement destination.

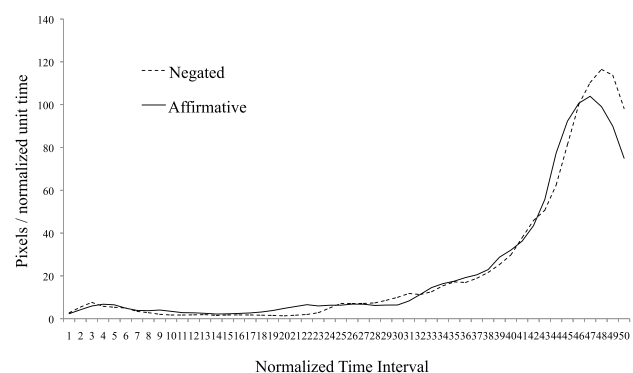


Figure 6: Instantaneous velocity by condition

Discussion

A main effect of negation was found where RT is higher for negated sentences. No differences were found in maximum deviations, and fine-grained spatial differences appear to be a function of different velocities. The averaged velocities exhibit a slowing and then greater acceleration for negated versions, indicating a hypothesis is being formed and information about possible alternatives is accumulating. The acceleration toward the end of the trajectories indicates that some activation has already accumulated for the possibilities, and so it is easier to fully activate the winning possibility faster because it already has some partial activation. This scenario can also explain previous results, where subjects respond faster to a picture of an eagle in the sky when the sentence is "The eagle is not in the sky", sky is one of the possibilities having partial activation, thus slightly primes responses to the affirmative version of that picture. The slowing down of the trajectory during the middle phase of the trajectory indicates the negating word acts as a contextual modifier, allowing for integration of immediate context (i.e. the two picture possibilities in this experiment), accounting for the longer amount of time it takes to process a negated sentence. This is also reflected in the maximum deviations: the fillers adhere closer to a straight trajectory toward the response, because there is no partial overlap of information, and thus the participant could begin moving confidently toward the response once the noun was heard. The lack of an effect between the affirmative and negated maximum deviations suggests that there is no direct competition between those two responses.

General Discussion

Negation may appear to require special mechanisms and extra explanations to account for its variability. However, the data presented here suggest that negation needs nothing more than some eyes and ears to be learned. The computational model created here demonstrates that learning negation as a contextual modifier is possible, and future explorations will aid in experimental validation. The results of the mousetracking experiment provide additional evidence that existing theoretical frameworks of incremental constraint-based processing can easily accommodate negated, as well as affirmative statements. Further, preliminary evidence suggests a transient period of increased competition in processing a negated statement, accounting for the increase in reaction times. Importantly, this must be a competition between various features the negated simulation could be composed of, and is *not* a competition between affirmative and negated representations.

Occam's razor dictates the most parsimonious explanation is the best explanation, which would be a one-stage processing requiring no additional interfacing processes. One major issue for future work will be that this has no predictions the logical model does not also have, thus leaving only null results as evidence to the contrary.

However, as is common knowledge, null results are not logically sound to rely upon, and so a defining property of the one-stage model could be sought after in the computation realm, to aid in discriminating these hypotheses.

Many aspects of the current model need to be examined in greater detail and tested experimentally. For example, the perceptual simulation in this model is not predictive, but because of word prediction is mildly emergently predictive. Because "heads up" is partially activated when "coin" is the input, "coin" ever so slightly activates some of the features of heads up. Another example is that this is a contextual modifier, and thus the negation may modulate a velocity profile as a function of location in a sentence. Despite the need for future work, the current research further supports that comprehending negation can be achieved through perceptual simulation rather than logical operators. Linguistic negation is not logical, but rather a process of integrating context with grounded, distributed meaning.

References

- Anderson, S.E., Huette, S., Matlock, T., & Spivey, M.J. (2010). On the temporal dynamics of negated perceptual simulations. In F. Parrill, V. Tobin, & M. Turner (Eds.), *Meaning, form, & body*, pp. 1-20. Stanford, CSLI.
- Cohen, J., MacWhinney, B., Flatt, M. & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers*, 25, 257-271.
- Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Howell, S. R., Jankowicz, D., & Becker, S. (2005). A Model of Grounded Language Acquisition: Sensorimotor Features Improve Lexical and Grammatical Learning. *Journal of Memory and Language*, 53, 258-276.
- Kaup, B., Yaxley, R.H., Madden, C.J., Zwaan, R.A., & Lüdtkke, J. (2006b). *Experiential Simulations of Negated Text Information. The Quarterly Journal of Experimental Psychology*, 60, 976-990.
- Samad, T. (1988). Toward connectionist rule-based systems. *IEEE International Conference on Neural Networks*, 2, 525-532.
- Spivey, M., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences of the USA*, 102, 10393-10398.
- Usher, M., & McClelland, J.L. (2003). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550-592.
- Wojnowicz, M.T., Ferguson, M.J., Dale, R., & Spivey, M.J. (2009). The self-organization of explicit attitudes. *Psychological Science*, 20, 1428-1435.