

Any more, this feature varies: An experimental study of non-polarity *any more*

Frances Blanchette & Valerie Keppenne*

Abstract. The non-polarity “positive” *any more* construction (NPAM; e.g., *I write letters any more*, ‘I write letters nowadays’) is a vernacular feature found in the Midland dialect region. Previous analyses have suggested that while its negative polarity item (NPI) counterpart (*I don’t write letters any more*) is associated with a positive presupposition and a negative assertion, NPAM contributes a negative presupposition and a positive assertion. This paper presents an experiment testing this hypothesis. The results provide initial support for the negative presupposition–positive assertion analysis of NPAM among speakers with regional exposure, and further suggest that such exposure is not necessary to know this feature does not share the same meaning as NPI-*any more*. We discuss implications of these results for our understanding of *any more* specifically and of polarity sensitivity more generally.

Keywords. polarity sensitivity; experimental semantics; regional variation; experimental syntax; Central Pennsylvania dialect

1. Introduction. “Positive” or non-polarity *any more* (NPAM; Horn 2021) is a vernacular feature found in the Midland region (Labov 1973; Wood et al. 2020). NPAM is said to have lost its polarity sensitivity since, unlike its NEGATIVE POLARITY ITEM (NPI) counterpart (1), it does not require a preceding negation.¹ This is shown in the different acceptability statuses of (1) and (2) when the negated auxiliary *don’t* is absent:

- (1) I *(don’t) write letters any more. (NPI)
- (2) I write letters any more. (NPAM)

In addition to the distinct polarity sensitivities of NPAM and NPI *any more*, constructions with these items also have clear semantic differences. While NPI *any more* can be understood to contribute a positive presupposition and a negative assertion (1a), NPAM has been hypothesized to reverse this pattern, contributing a positive assertion and a negative presupposition (2a) (Hindle & Sag 1975; Horn 2021, see Kuhn & Maldonado 2024 for a recent review):²

- (1a) I don’t write letters any more. (NPI)

Positive presupposition: I used to write letters.

Negative assertion: I do not write letters nowadays.

* This research was conducted with funds from the Penn State Eberly College of Science, for which we are very grateful. We are indebted to the REEDS (Re-examining Dialect Syntax) Network for the opportunity to present and discuss an earlier version of this work. We are also grateful to Larry Horn for helpful feedback on our experiment design, and to Samantha Roth and Elle Tang for help analyzing the data. Thanks also to the attendees at our LSA talk for helpful comments and feedback, especially Larry Horn and Ezra Keshet. All mistakes are ours, of course. Authors: Frances Blanchette, Penn State (fkb1@psu.edu) & Valerie Keppenne, Penn State (vxk57@psu.edu).

¹ While negation is a prototypical environment for NPI “licensing”, the conditions are more general and can be characterized as downward entailing (Ladusaw 1979) or non-veridical (Giannakidou 1998).

² Larry Horn and Ezra Keshet (p.c.) note that the apparent negative presupposition of NPAM may not in fact be a presupposition but rather an implicature, and Larry Horn further suggests that the implicature of NPAM may simply be that a change has occurred. We are currently working on testing these possibilities, but for now it suffices to state that the methods and results we report in this paper are compatible with alternative analyses such as these.

(2a) I write letters anymore. (NPAM)

Negative presupposition: I did not used to write letters.

Positive assertion: I write letters nowadays.

While much has been learned about the distribution of NPAM from both corpus studies (e.g., Strelluf 2019 and references therein) and surveys (e.g., Wood et al. 2020) (see Maher & McCoy 2011 for a review), many questions remain. This paper investigates one such question, namely, whether we can find experimental support that speakers assign something like a negative presupposition–positive assertion interpretation to NPAM constructions.

Since we are interested in speakers’ intended meaning when they use this construction, and since it is difficult to infer speaker intent from corpus data, it is necessary to probe intuitions a bit more directly to draw inferences about the semantics of NPAM. However, since NPAM is a regional, non-mainstream form and therefore subject to much variation, we cannot assume that any given individual from a region where NPAM has been identified will have robust knowledge of this feature. In addition, since the level of meaning we are interested in is relatively nuanced, a single observation from a given individual may not suffice, and it may be necessary to obtain repeated measures from multiple individuals in order to be able to draw stronger and more direct inferences about interpretation.

We therefore designed an experiment to tap into people’s interpretations of NPAM, and whether the hypothesized interpretation of a negative presupposition–positive assertion correlates with a readiness to accept this construction. Our experiment consisted of two tasks: a meaning judgment task designed to capture English speakers’ intuitions about the meaning of NPAM and its relationship to NPI *anymore*, and a naturalness rating task that aimed to explore their intuitions about the relative naturalness of these constructions. As a preview to our results, we find some initial experimental support that NPAM does indeed reverse the meaning of its NPI counterpart, contributing a negative presupposition and a positive assertion instead of positive presupposition and a negative assertion. Adding to this, we find that while regional exposure appears to be necessary for this level of knowledge to be displayed, it is not necessary to know that NPAM does not contribute the same meaning as an NPI construction with *anymore*. This novel finding emerged from our experimental design, which included a control group of participants from regions where NPAM is not in regular use. Our results therefore afford new insight into how speakers first encountering this construction respond to it, and the possibilities and limitations afforded by their initial representations.

2. The meaning judgment task. We designed a meaning judgment task to test whether NPAM is associated with a negative presupposition and a positive assertion reading.

2.1. MEANING JUDGMENT TASK DESIGN. We adopted the covered picture task paradigm in Bill et al. (2018), which tested similar types of semantic knowledge with more mainstream features. Participants were shown two sets of images, one covered and one uncovered, with a sentence below them (Figure 1). The uncovered picture set had three identical images with three sequential years written above them, ending in the current year. This was intended to evoke the meanings of now (the current year) and before (the two previous years). Large red Xs were used to represent the negative component of the meaning: Xs on the first two pictures were intended to elicit the meaning of ‘not before’, while if the last picture (the current year) was Xed out this was intended to elicit a ‘not now’ reading. Participants were asked to select the set of pictures, covered or uncovered, that best represented a given sentence’s meaning. Selection of the set of covered

pictures was therefore essentially a rejection of the meaning depicted by the uncovered set of pictures.

Figure 1 contains an example of a stimulus item where the hypothesized NPAM negative presupposition–positive assertion reading matches the uncovered picture. If participants systematically selected the uncovered picture for items like the one in Figure 1, we took this as support for the hypothesis that they interpreted NPAM as contributing a negative presupposition and a positive assertion. If, on the other hand, they selected the covered picture set for items like this one, we took this as evidence that they did not generate the negative presupposition–positive assertion reading of NPAM constructions.

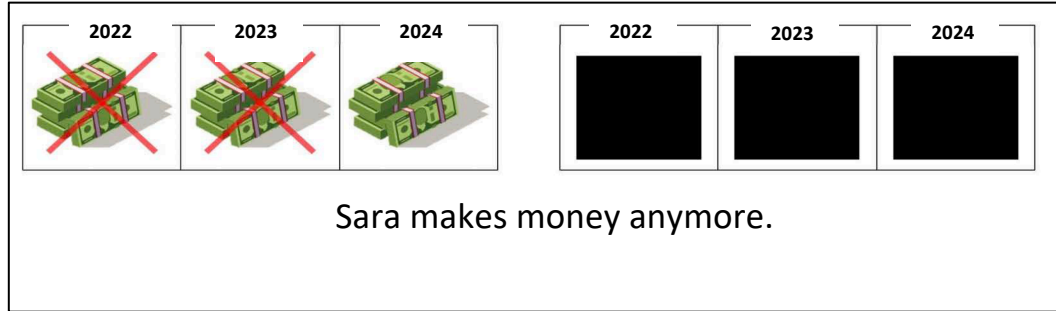


Figure 1. NPAM-match stimulus example from the meaning judgment task

The meaning judgment task contained 32 experimental items split across four Latin Square lists, as illustrated in Table 1 below. Each individual item or set of images occurred in each of four conditions: (i) as an NPI construction where the uncovered picture matched the positive presupposition–negative assertion reading characteristic of NPI constructions (NPI-match), (ii) as an NPI construction where the uncovered picture did not match the characteristic NPI interpretation (NPI-mismatch), (iii) as an NPAM construction where the uncovered picture matched the hypothesized positive assertion–negative presupposition interpretation (NPAM-match), and (iv) as an NPAM construction where the uncovered picture did not match the hypothesized interpretation (NPAM-mismatch). Participants were randomly assigned to one of the four lists such that, while each individual participant saw 32 experimental items in total – 8 NPI-match, 8 NPI-mismatch, 8 NPAM-match, and 8 NPAM-mismatch – they only saw one of the four versions of each item. In addition to the 32 experimental items, participants also judged the meaning of 96 fillers of similar complexity, leading to a total of 128 items.













List 1	List 2	List 3	List 4
Sara doesn't make money anymore.	Sara doesn't make money anymore.	Sara makes money anymore.	Sara makes money anymore.
  	  	  	  
NPI-match (select picture)	NPI-mismatch (select covered picture)	NPAM-match (select picture)	NPAM-mismatch (select covered picture)

Table 1. Counterbalanced design for the meaning judgment task

As Table 1 shows, for the match and the mismatch items, the expected or (hypothetically) “correct” answer was effectively reversed for the NPI and NPAM items. For example, in Table 1, the picture with Xs on the two previous years is a mismatch for an NPI sentence (List 2), so if

participants are seeing that picture in conjunction with an NPI sentence, then the correct answer would be to select the uncovered picture. That same picture, however, would be a match for an NPAM sentence (List 3) under the hypothesized interpretation. This manipulation will be important in our results and discussion below.

2.2. EXPERIMENT PARTICIPANTS. We recruited 46 participants from Central Pennsylvania (henceforth the PA group), and 57 participants from Amazon Mechanical Turk. For the Mechanical Turk participants, a prerequisite for participation was that they be from areas that do not overlap with the Midland Region (henceforth the non-Midland group), since based on previous research people from these regions are likely to have had limited exposure to NPAM (Wood et al. 2020). We excluded Mechanical Turk participants who, despite this prerequisite, reported to being from the Midland region in a post-task survey, as well as participants who in the post-task survey reported to having regular exposure to NPAM. We also excluded participants who on the basis of their performance on filler and control trials did not appear to understand the task. This left us with 46 participants in the non-Midland group, and 46 from the PA group since no exclusions were necessary there. Participants in both groups were of similar ages, with the PA group ranging from 19-66 (mean 39) and the non-Midland group ranging from 23-75 (mean 39).

In terms of gender identity, the PA group was somewhat skewed toward female, with 36 participants identifying as female and 10 as male. The non-Midland group was more balanced between male (24) and female (21), with one participant identifying as non-binary. Since gender is known to shape linguistic variation (e.g., Eckert 1989), it is possible that the imbalance in the PA group may have skewed our results somewhat. For example, since we are interested in participants' behaviors related to a vernacular, non-mainstream variant, and since some studies have shown rates of vernacular feature use to be lower in women's than in men's speech, the extent of NPAM knowledge in the greater PA population may be underrepresented in our data set. We therefore interpret our data with this potential limitation in mind, though we do not think it merits a central role in our results interpretation.

In terms of racial and ethnic identity in the PA group, one participant identified as African American, Hispanic, and White; one identified as Asian American, one as Indian, one as White and Asian, and one as White and Korean, while the remaining 41 identified as White or Caucasian. In the non-Midland group, five participants identified as Asian, three identified as Black, one identified as White and Latino, and one identified as White and Native American, while the remaining 36 identified as White or Caucasian. Participants were therefore similar across groups in terms of their racial and ethnic identities, with the majority identifying as White or Caucasian.

All participants from the PA group reported to growing up in and spending the majority of their lives in a range of counties within the Central PA region. Participants in the non-Midland group were from a wide range of US states including Alabama (1), Arizona (1), California (6), Florida (4), Georgia (2), Kentucky (1), Michigan (6), New Hampshire (2), New York (2), Illinois, Kentucky (2), Maryland (1), New Jersey (1), North Carolina (1), Oregon (1), Rhode Island (1), Texas (5), Tennessee (3), Virginia (2), Washington (2), and Wisconsin (2).

2.3 MEANING JUDGMENT TASK RESULTS. Figure 2 shows the results of the meaning judgment task.³ As the figure shows, both groups had high accuracy for both the NPI match and the NPI

³ Results from five of the non-Midland participants who met our demographic inclusion criteria but did not appear to understand the task based on their responses to some of the control and filler items were excluded from analyses of the meaning judgment task. The graphs and statistical modeling results reported for the meaning judgment task are therefore reflective of 41 participants as opposed to the 46 whose demographic results we reported above. These

mismatch trials. This means that, in general, participants were able to identify NPI constructions as compatible with a positive presupposition and a negative assertion, and reject the reverse (i.e., the NPAM) meaning for these constructions. This makes sense, since NPI constructions are a mainstream feature and occur in virtually all English varieties, both vernacular and standardized (see, e.g., Blanchette 2015). NPAM trials, on the other hand were much more variable. Both groups had relatively low accuracy on the NPAM match trials, but the PA group had slightly higher accuracy overall than the non-Midland group. This represents the first experimental evidence that being from Pennsylvania makes you more likely to interpret NPAM as having a negative presupposition–positive assertion reading. Interestingly, accuracy was much higher on the NPAM mismatch trials, suggesting that both groups found it relatively easy to reject the idea that these have the same meaning as NPI constructions. We return to this in the discussion.

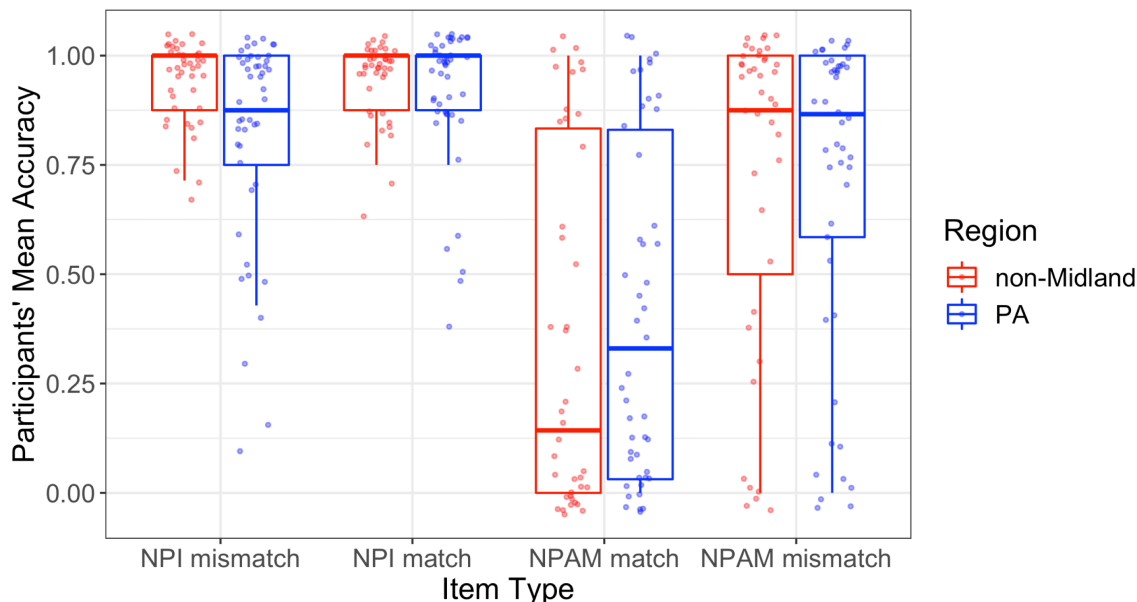


Figure 2. Accuracy on the meaning judgment task, with box plots showing overall quartiles and median, and points illustrating individual participants’ mean accuracy levels. Values for non-Midland participants are shown in red, and for PA participants in blue.

To better understand these results, we fit a general linear mixed effects regression to the data using the *glmer()* function of the *lme4* package (version 1.1.29, Bates, Maechler, Bolker & Walker 2015) in R (version 4.2.0; R Core Team 2022). Predictor variables, entered into the model using zero-centered effects coding, were the within-participants factors *anymore* type (NPAM, contrast code: -0.50 vs. NPI, contrast code: 0.50) and trial type (match, -0.50 vs. mismatch, 0.50), and the between-participants factor region (non-Midland, contrast code: -0.50 vs. PA, contrast code: 0.50). We included an interaction term for *anymore* type and region to explore differences in interpretation across groups. We also included random intercepts for participant and item and random slopes for *anymore* type and trial type by participant.

Model results are shown in Table 2. The main effect of trial type was reliable because match trials were more difficult for participants overall than mismatch trials ($Pr(>|z|) < .001$). Though

five excluded participants’ responses were included in the analyses of the ratings task, since their ratings of control items suggested they did understand the task, but excluded from the comparison of participants’ behaviors across tasks.

inconsistent with previous research using this paradigm (Bill et al., 2018), this effect appears to have been driven by the NPAM trials, since participants tended to have much higher accuracy with NPAM mismatch trials than with NPAM match trials (which we discuss further below). The main effect of *anymore* type was reliable because NPI items were more likely to be answered correctly than NPAM items ($Pr(>|z|) < .001$). There was no reliable effect of region, since overall non-Midland and PA participants had similar accuracy levels across all trial types ($Pr(>|z|) = .12$). However, there was a marginally reliable interaction between *anymore* type and region ($Pr(>|z|) = .07$), suggesting that participants from PA did not treat NPI and NPAM items in exactly the same way.

Effect	Estimate	Std. error	z-value	Pr(> z)
(Intercept)	2.02	0.17	11.62	< .001
Trial Type (mismatch–match)	1.92	0.31	6.18	< .001
<i>anymore</i> type (NPI–NPAM)	3.10	0.31	10.07	< .001
Region (PA–non-Midland)	-0.40	0.31	-1.30	.12
<i>anymore</i> type x Region	-1.05	0.58	-1.82	.07

Table 2. Model results for the meaning judgment task

Planned comparisons further explored the interaction between *anymore* type and region. The focus here was on the NPAM match and mismatch trials only, since unlike the NPI construction, NPAM is regional and PA speakers are predicted to have more knowledge of it than non-Midland speakers. Two models, identical to the main model except for their contrast codes, were fitted to the data to examine the simple main effect of region within the NPAM mismatch and the NPAM match trials. These models revealed no reliable difference between non-Midland and PA participants in the NPAM mismatch trials (Estimate = 0.09, $se = 0.48$, $z = 0.187$, $Pr(>|z|) = .09$), but a reliable difference between regions in NPAM match trials (Estimate = -0.92, $se = 0.36$, $z = -2.581$, $Pr(>|z|) < .01$), reflecting a small but meaningful regional difference in NPAM interpretation between groups in the predicted direction. The lack of meaningful difference between groups in the NPAM mismatch trials reflects the intriguing fact that, in contrast with the match trials, both groups had similarly high accuracy levels for these. We discuss possible reasons for this further below.

3. The naturalness rating task. Participants also completed a rating task immediately following the meaning judgment task.

3.1. RATING TASK DESIGN. The stimuli for the rating task consisted of items constructed for the meaning judgment task – both NPI and NPAM, as well naturalistic NPAM items adapted from Strelluf’s (2019) Twitter corpus study.⁴ Since each item appeared as either an NPI or an NPAM construction on the meaning judgment task, these items were split across two lists, such that participants were asked to rate only the version that they did not see in the previous task, to avoid repetition of identical stimuli. Each participant rated 16 NPI and 16 NPAM items constructed for the meaning judgment task and 8 naturalistic NPAM items on a scale of 1-7, where one is

⁴ Strelluf observes a wide range of different “trigger” types (e.g., inherent negative) for the tokens of *anymore* in his corpus, as well as instances of NPAM with “no trigger”. We adapted only tokens with no trigger, since these were the most semantically parallel to the items from our meaning judgment task. We return to this in the discussion section.

completely unnatural and 7 is completely natural. They also rated 24 fillers of similar complexity, for a total of 44 items. Examples of the critical items are shown in Table 3.

Item Type	Example
16 NPI items from meaning judgment task	<i>Sara doesn't enjoy music anymore.</i>
16 NPAM items from the meaning judgment task	<i>Alex takes pictures anymore.</i>
8 naturalistic items adapted from Strelluf (2019) (NPAM-nat)	<i>That player is playing how we expect anymore.</i>

Table 3. Example items on the naturalness rating task

3.2. RATING TASK RESULTS. Figure 3 illustrates the results of the naturalness rating task. As the figure shows, both groups rated NPIs highly and gave low ratings for NPAM, including both the items created for the meaning judgment task as well as the more naturalistic ones. Moreover, despite the fact that the PA group displayed greater systematic interpretation of NPAM as contributing a negative presupposition and a positive assertion on the meaning judgment task, their ratings of the NPAM sentences that appeared on the meaning judgment task appear identical to those of the non-Midland group, who did not display systematic understanding of the construction's meaning. However, while ratings for the naturalistic NPAM items are also low, there do appear to be some group differences for those items in the predicted direction: as a group the PA participants rated these more highly than their non-Midland counterparts, a fact which we can attribute to regional exposure to the construction.

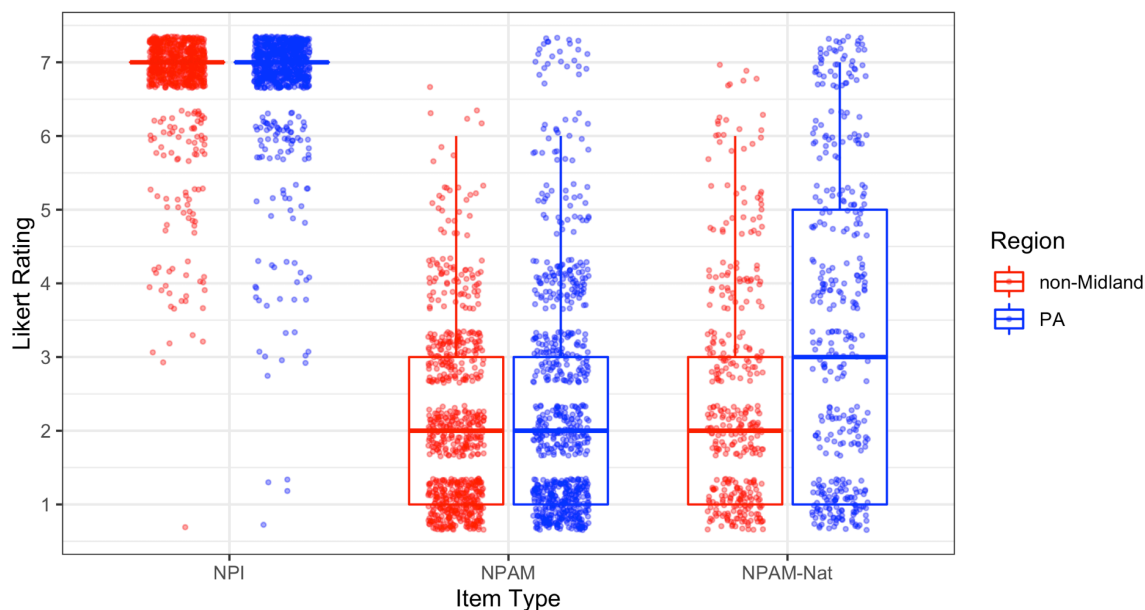


Figure 3. Results of the naturalness rating task, with box plots showing overall quartiles and median and points showing ratings for a single item by a single participant. Values for non-Midland participants are shown in red (left), and for PA participants in blue (right).

To better understand these results, we fitted a cumulative link mixed model to the data using the *clmm()* function of the *ordinal* package (version 2019.4-2; Christiansen 2019) in R, with a probit link function. Since exploratory data analysis revealed that the groups treated the NPI items nearly identically (as seen in Figure 3 above), these were not included in the model. Predictor

variables, again entered using zero-centered effects coding, were the within-participants factor NPAM type (NPAM-Experimental, contrast code: -0.50 vs. NPAM-Naturalistic, contrast code: 0.50) and the between-participants factor region (non-Midland, contrast code: -0.50 vs. PA, contrast code: 0.50). An interaction term for NPAM type and region was also included to explore whether judgments were different between groups. We included random intercepts for participant and item and random slopes for NPAM type by participant and region by item.

Model results are shown in Table 4. The main effect of NPAM type was reliable, since naturalistic NPAM items were rated more highly overall than the NPAM items constructed for the meaning judgment task ($Pr(>|z|) < .001$). There was a marginal effect of region, since overall PA participants gave slightly higher ratings than non-Midland participants ($Pr(>|z|) = .09$). There was also a small but reliable interaction between NPAM type and region ($Pr(>|z|) < .05$), suggesting regional differences in the extent to which participants distinguished between NPAM items from the meaning judgment task and more naturalistic NPAM items adapted from Twitter.

Effect	Estimate	Std. error	z-value	Pr(> z)
NPAM Type (NPAM-Nat–NPAM-Experiment)	1.19	0.21	5.68	< .001
Region (PA–non-Midland)	0.79	0.46	1.72	.09
NPAM type x Region	0.58	0.29	1.99	< .05

Table 4. Model Results for the Naturalness Rating Task.

This primary model was again followed by further analyses to explore the interaction. These models revealed no reliable difference in ratings between regions for the NPAM-experiment items (Estimate = 0.49, $se = 0.53$, $z = 0.93$, $Pr(>|z|) = .35$), but a reliable difference between regions for the NPAM-naturalistic items (Estimate = 1.08, $se = 0.42$, $z = 2.57$, $Pr(>|z|) < .01$). This indicated that, although ratings of the NPAM items that appeared on the meaning judgment task were similarly low across the PA and the non-Midland groups, PA participants systematically rated the more naturalistic NPAM items higher than participants not from the Midland dialect region. This supports the conclusion that being from Pennsylvania makes a speaker more likely to accept NPAM than if they were from outside the Midland region.

4. Connecting meaning judgments and naturalness ratings. In addition to analyzing the meaning judgment and naturalness rating results independently of one another, we also explored relationships between participants’ behaviors related to NPAM on the two tasks. We were particularly interested in whether accuracy on the NPAM match items, where participants explicitly selected the negative presupposition–positive assertion reading for NPAM. This was because these items provided the most clear and transparent view of participants’ NPAM interpretation. If participants who tended to select the predicted negative presupposition–positive assertion reading of this construction also tended to rate NPAM items highly, then we reasoned that this could be taken as further support for the hypothesized interpretation of NPAM as having the reverse meaning of NPI *anymore* under negation.

Figure 4 illustrates the relationship between participants’ accuracy on the NPAM-match items in the meaning judgment task and their ratings of the two types of NPAM item included on the naturalness rating task. Ratings of NPAM items constructed for the experiment are shown in the left panel, and ratings of more naturalistic NPAM items are on the right.

As Figure 4 shows, the PA group’s ratings of both types of NPAM items increase as a function of accuracy on the meaning judgment task. This suggests that a PA speaker’s interpretation

of an NPAM item as a positive assertion with a negative presupposition is systematically and positively related to that speaker’s acceptance of NPAM constructions as a natural way of speaking. However, importantly, the same is not the case for the non-Midland group, since better performance on the meaning judgment task does not relate systematically to higher ratings of NPAM items. Since non-Midland speakers did not show the same pattern, this relationship between acceptability and interpretation as a positive assertion–negative presupposition appears to be tied to regional knowledge.

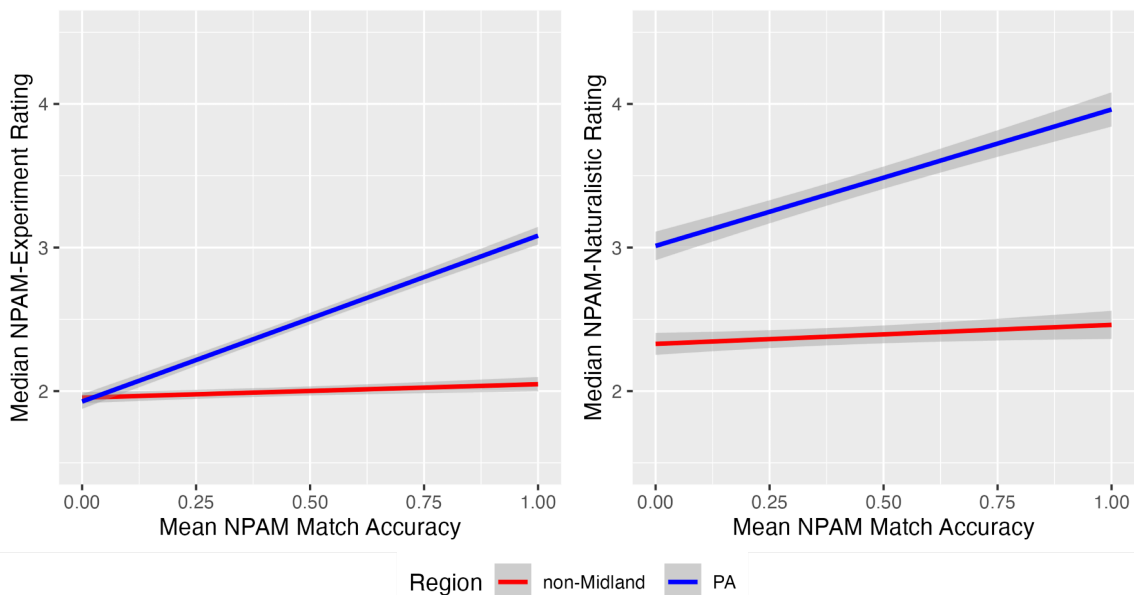


Figure 4. Accuracy on the NPAM match items on the meaning judgment task as a function of median acceptability of two types of NPAM items in the naturalness rating task. Lines were generated with the “geom_smooth” function (method “glm”) in R.

5. Conclusions and discussion. Since NPAM is a regional feature that is not considered part of the “mainstream” or standardized variety, it is not surprising that the experiment results revealed substantial amounts of interspeaker variation in both the meaning judgments and naturalness ratings of this construction. Nevertheless, the experimental methods employed in the study allowed us to observe some patterns. We summarize these here and discuss what they imply for representations of NPAM and polarity sensitivity more generally.

5.1. EXPERIMENTAL SUPPORT FOR A REGIONAL INTERPRETATION. One conclusion that emerges from our results is that being from Pennsylvania makes you more likely to interpret NPAM as contributing a negative presupposition and a positive assertion than if you are from a region outside the Midland. This conclusion provides the first experimental support for existing hypotheses about NPAM meaning (Hindle & Sag 1975, Horn 2021), and further illustrates that the ability to generate the negative presupposition–positive assertion reading of NPAM is dependent upon regional exposure. We found that while participants from Pennsylvania and those from non-Midland areas readily associated NPI constructions with a positive presupposition–negative assertion reading, only the PA participants reliably selected the reverse negative presupposition–positive assertion reading for sentences with NPAM. While the pattern was not categorical for PA

participants, the differences between groups were statistically reliable and pointed toward meaningful regional differences in NPAM interpretation.

5.2. SUPRA-REGIONAL DIFFERENCES IN NPAM AND NPI INTERPRETATION. An additional conclusion that emerges from our results is that regional exposure is not necessary to know NPAM and NPI constructions have different meanings. This was borne out in the high accuracy rates for the NPAM mismatch trials in the meaning judgment task. In these trials, participants were judging an NPAM sentence (e.g., *Sara makes money anymore*) along with a covered picture and an uncovered picture that reflected a meaning consistent with the NPI interpretation of that sentence (e.g., *Sara doesn't make money anymore*). Participants from both groups, noticing that the negation was absent from the NPAM sentence, readily rejected the NPI positive presupposition–negative assertion meaning for NPAM, selecting the covered picture instead. The fact that the non-Midland group performed just as well as the PA group in this condition illustrates that regional exposure is not necessary to know that NPAM *anymore* does not contribute the same meaning as NPI *anymore*.

5.3. REGIONAL SIMILARITIES AND DIFFERENCES IN NATURALNESS RATINGS. Regional differences in behaviors toward NPAM were also apparent in the naturalness rating task, but not across the board. In particular, we saw that when rating items constructed for the meaning judgment task, regional exposure did not seem to matter for our participants, since both groups gave NPAM sentences constructed for the experiment relatively low ratings. Crucially, this was due to NPAM itself and not some property of the items more generally, since the same items with NPI *anymore* were given high ratings by both groups. This fact reveals an asymmetry between acceptability and interpretation of NPAM, since the PA group did in fact demonstrate greater comprehension of the lab-constructed items in the meaning judgment task. The fact that the PA group nevertheless did not find these same lab-constructed NPAM items more acceptable than the non-Midland group demonstrates an asymmetry between acceptability and interpretation. Thus, while NPAM is attested and reliably associated with a predictable meaning in Central and Western Pennsylvania, it has not yet reached the level of NPI *anymore* in terms of its ability to appear across registers and in a range of different situational contexts, including laboratory experiments (see also Labov 1973).

Although sentences were not rated differently across groups for the lab-constructed NPAM items, regional differences did emerge reliably in ratings of the naturalistic NPAM items adapted from Strelluf's (2019) corpus study. Strelluf reported on a wide range of non-negative contexts for *anymore*, but we adapted our items only from ones he labeled as having clearly non-negative meaning ("no trigger", in his terms; "veridical" in Giannakidou's (1998) terms). Rates of these types of NPAM constructions were relatively low in Strelluf's data, collected from Tweets across the Midland Region. However, the western Pennsylvania city of Pittsburgh, which is adjacent to the region in which our data were collected, was found to have the highest index for these clearly positive instances of NPAM (Strelluf 2019: 334, Fig. 3).

In our experiment, for these more naturalistic NPAM items, while the ratings were still well below the median ratings for NPI constructions, the PA group reliably rated them more highly than the non-Midland group. This illustrates that, despite their ability to reliably assign a predictable meaning to lab-constructed NPAM, Pennsylvania speakers prefer to encounter this regional feature embedded in more naturalistic linguistic contexts such that they may find in everyday conversation. This finding supports the conclusion that our PA participants have regional knowledge of NPAM, since in addition to the meaning judgment task, group differences also

emerged in the naturalness rating task. The fact that the PA group distinguished between the lab-constructed and the more naturalistic NPAM items while consistently rating the lab-constructed NPI items very highly further highlights the distinct social statuses of NPI and NPAM constructions, and the vernacular, non-standardized status of NPAM.

5.4. THE IMPORTANCE OF RELATING MEANING TO NATURALNESS. Perhaps the clearest support for the negative presupposition–positive assertion interpretation of NPAM was found in the relationship (or lack thereof) between participants’ behaviors across the two tasks. For the PA group only, ratings increased as a function of accuracy on the NPAM match trials. This means that the more a PA participant systematically selected a picture that matched the negative presupposition–positive assertion interpretation of NPAM, the more likely they were to give higher naturalness ratings of NPAM. This was true for both the lab-constructed and the more naturalistic items. Since NPAM is a variable feature that likely occurs with varying degrees of frequency in our PA participants’ input, this connection between meaning judgments and naturalness ratings is important. If we take higher ratings to be indicative of an individual’s exposure and usage of a construction, and if we assume exposure and usage lead to stronger grammatical representations within a speech community, then this systematic relationship between meaning and naturalness suggests that the construction is indeed associated with something like a negative presupposition–positive assertion reading in communities with NPAM such as those in Central and Western Pennsylvania.

5.5. QUESTIONS AND IMPLICATIONS. While the study results reported here contribute some new information about the meaning and regional, vernacular status of NPAM, as with any complex topic such as this, a single study is far from sufficient, and many questions remain. One set of questions pertains to the relationship between NPAM and NPI *anymore*, and more generally, to the relationship between “positive” and “negative” uses of items such as *anymore*, *yet*, and others more standardly found in NPI environments. What is the appropriate way to characterize speakers’ grammatical knowledge of these pairs, which are similar in form but completely different in polarity sensitivity? And what kinds of connections are speakers making in their minds, if any, between these different constructions upon encountering them?

For at least a subgroup of the PA speakers who participated in our study, there is a clear and systematic semantic relationship between positive and negative uses of *anymore*: semantically, NPAM reverses the assertional and presuppositional content of a parallel NPI construction with this term. But what about the non-Midland speakers, most if not all of whom were encountering NPAM for the first time in this experiment? And what about those PA participants with lower accuracy on the meaning judgment task and correspondingly lower ratings on the naturalness rating task? When encountering NPAM paired with an image that depicted an NPI *anymore* reading, these participants were quick to reject this reading. Importantly, however, they were not nearly as quick to select the meaning for NPAM that reversed the presuppositional and assertive content of the NPI construction.

What, then, were participants who did not systematically select the negative presupposition–positive assertion reading for NPAM doing with this construction? While our experiment results leave this as both a figurative and a literal black box (given the option to select a covered picture when they did not want to choose the visible one), we can speculate to some extent at this stage. It was clear in our results that, regardless of group membership, participants knew that NPI *anymore* contributed a negative (or negated) assertion. In addition, in contexts where *anymore* was not preceded by a negation, they seemed to know that the negative (or negated) assertion

interpretation was no longer possible. However, this knowledge did not readily extend itself to a reversal of truth conditions, since those without regional knowledge (and even many within the NPAM region) were nevertheless reluctant to select the positive assertion interpretation.

If we assume that NPAM and NPI *anymore* are underlyingly related, it is possible that those participants who rejected both the positive and the negative assertion readings for NPAM were entertaining some intermediate stage of representation for *anymore*. For instance, they may have been reluctant to allow that *anymore* could lack polarity sensitivity, but nevertheless concluded that on its own it could not contribute a semantic negation. Non-negative uses of polar-sensitive *anymore* are well attested in mainstream and vernacular speech, and are instantiated in a range of construction types such as polar questions (e.g., *Does Sara make money anymore?*). While varying analyses of NPIs in these types of non-negative environments exist (compare, for example, Giannakidou 1998 vs. Collins and Postal 2014), they are generally understood to reflect polarity sensitivity. In polar questions, for example, it is the polar interrogative that provides the appropriate “licensing” conditions.

While polar questions such as *Does Sara make money anymore?* do not assert anything, they do appear to contribute the same presuppositional content as parallel constructions with an NPI under negation. For example, both *Does Sara make money anymore?* and *Sara doesn't make money anymore* presuppose that Sara used to make money. Participants who did not select the negative presupposition–positive assertion interpretation for NPAM but nevertheless rejected the parallel NPI under negation reading may have been attempting to map NPAM onto their existing representations of non-negative NPIs. Since accuracy on the NPAM trials in this experiment required the selection of a negative and not a positive presupposition, this presuppositional difference in interpretation between NPAM and non-negative NPI constructions could be at the heart of participants' differing behaviors toward NPAM. In particular, it may be that a certain amount of regional exposure is required to know that positive (or at least non-negative) uses of *anymore* may contribute a negative presupposition.

References

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1–48.
<https://doi.org/10.18637/jss.v067.i01>.
- Bill, Cory, Jacopo Romoli & Florian Schwarz. 2018. Processing presuppositions and implicatures: similarities and differences. *Frontiers in Communication* 3. 44.
<https://doi.org/10.3389/fcomm.2018.00044>.
- Blanchette, Frances. 2015. English negative concord, negative polarity, and double negation. New York, NY: CUNY Graduate Center dissertation. <https://ling.auf.net/lingbuzz/002654>.
- Christensen, Rune Haubo B. 2019. ordinal – Regression models for ordinal data. R package version 2019.4-25. <http://www.cran.r-project.org/package=ordinal/>.
- Collins, Chris & Paul Postal. 2014. *Classical NEG Raising: An Essay on the syntax of negation*. Cambridge, MA: MIT Press.
- Eckert, Penelope. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1(3). 245–267. <https://doi.org/10.1017/S095439450000017X>.
- Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins.

- Hindle, Donald & Ivan Sag. 1975. Some more on *anymore*. In Ralph W. Fasold & Roger Shuy (eds.), *Analyzing variation in language: Papers from the Second Colloquium on New Ways of Analyzing Variation*, 89–11. Washington, DC: Georgetown University Press.
- Horn, Laurence R. 2021. Anymore once more: Geographical and syntactic distribution. *Yale Working Papers in Grammatical Diversity*. <https://elischolar.library.yale.edu/ygdp/10/>.
- Kuhn, Jeremy & Mora Maldonado. 2024. The many meanings of an NPI: The view from artificial language learning. *PsyArXiv* manuscript. <https://doi.org/10.31234/osf.io/cpg7b>.
- Labov, William. 1973. Where do grammars stop? In Roger Shuy (ed.), *Report of the Twenty-Third Annual Round Table Meeting on Linguistics and Language Studies*, 43–88. Washington, DC: Georgetown University Press.
- Ladusaw, William A. 1979. *Polarity sensitivity as inherent scope relations*. Austin, TX: University of Texas dissertation.
- Maher, Zach & Tom McCoy. 2011. Positive *anymore*. Yale grammatical diversity project: English in North America. Website. <https://ygdp.yale.edu/phenomena/positive-anymore>.
- R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Strelluf, Christopher. 2019. *Anymore*, it's on Twitter: Positive *anymore*, American regional dialects, and polarity licensing in tweets. *American Speech* 94. 313–351. <https://doi.org/10.1215/00031283-7587883>.
- Wood, Jim, Kaija Gahm, Ian Neidel, Sasha Lioutikova, Luke Lindemann, Lydia Lee & Josephine Holubkov. 2020. Mapbook of syntactic variation in American English: Survey results 2015–2019. *Yale Working Papers in Grammatical Diversity*. <https://elischolar.library.yale.edu/ygdp/7/>.