Are Infants Sensitive to Informant Reliability in Word Learning?

Alayo Tripp, Naomi Feldman, and William Idsardi

1. Introduction

Common wisdom among adults when listening to speech is to "consider the source." The identity of a speaker can provide a wealth of context in interpreting a speech act. Despite this, social reasoning about differences between sources has not heretofore been considered part of early word learning, and is not traditionally included in models of infant word learning (e.g. Pinker, 1979; Chater and Manning, 2006; Frank et al., 2009). Instead, these models assume learners trust data from all informants equally. However, there is significant evidence from non-linguistic domains that infants use epistemic trust (Zmyj et al., 2010; Poulin-Dubois et al., 2011; Tummeltshammer et al., 2014).

Epistemic trust, the process of evaluating information sources for reliability, is robustly attested in the language processing of older children, who exhibit a clear bi-directional relationship between linguistic inference and inferences about reliability. In forced choice tasks, children's perception of informant reliability is influenced by observation of labeling behavior, and their endorsement of object labels is influenced by the familiarity and reliability of the adults providing them (Koenig and Harris, 2005; Corriveau and Harris, 2009). There is also extensive evidence that infants' linguistic perception predicts their attention to sources in non-linguistic tasks (Kinzler et al., 2005). If the ability to discriminate between speech sources for reliability is present earlier in life, it may influence infant behavior on word learning tasks.

One way to investigate the question of whether social inference influences early word learning is to begin with the existing empirical literature. A large body of work has used an audiovisual habituation experiment called the Switch Task to investigate early word learning. Results from Switch task experiments show that infants who can perform well on a task discriminating two lexical neighbors, or words which differ by a single phoneme (e.g., "buk" and "puk"), nevertheless do not consistently discriminate those same labels after being habituated to the presentation of these speech tokens as the labels of two different objects (Stager and Werker, 1997). This difficulty does not appear for pairs of words which differ by multiple phonemes (e.g., "lif" and "neem") (Werker et al., 1998). Slightly

* Corresponding author: Alayo Tripp, University of Minnesota, tripp158@umn.edu.
This research was supported by NSF grants BCS-1320410 and IIS-1421695.

© 2020 Alayo Tripp, Naomi Feldman, and William Idsardi. *Proceedings of the 44th Boston University Conference on Language Development*, ed. Megan M. Brown and Alexandra Kohut, 662-676. Somerville, MA: Cascadilla Press.

older infants show significantly improved performance, with 17-month-old infants being successful at learning the phonetically similar words (Werker et al., 2002). If changes in infant behavior on these tasks can be attributed to a developing ability to rely on inferences about informant reliability when interpreting language, this would provide evidence that socially informed reasoning about the selection of linguistic sources guides infants' early word learning.

In this paper we focus on one variation on the Switch task experiment by Rost and McMurray (2009), which showed that exposure to multiple speakers during habituation helps support 14-month-old infants' success on the Switch task. We apply a model adapted from Shafto et al. (2012) to simulate the experiments in Rost and McMurray (2009), and show that the source selection account can explain their results. We conclude by discussing the source evaluation hypothesis presented here in light of other variations on the Switch task, and implications for understanding the developmental trajectory of the lexicon.

2. The Switch Task

The Switch task audiovisual habituation procedure is a test of word learning wherein infants are presented with a repeated word paired with a visual display of a novel object. Infants observe an object and hear it labeled with a novel word until their looking time drops below a preset level, the habituation criterion. Then the infants' ability to discriminate the presented words is assessed using two types of trials. On *same* trials, the participants are exposed to the same object-word pairing(s) seen during habituation. On *switch* trials, they again see one of the familiarized objects, but this time it is paired with a mismatched label.

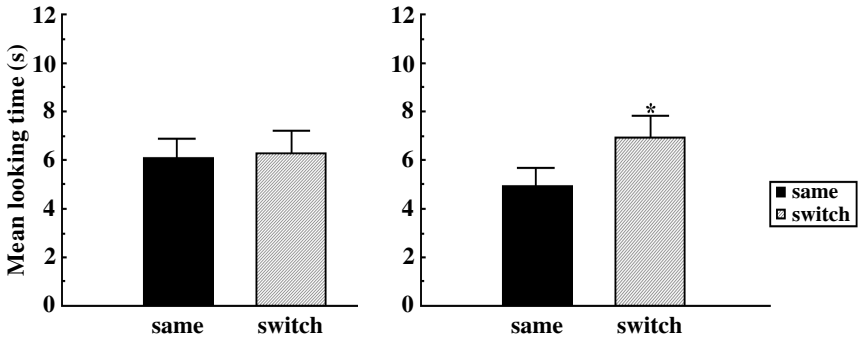
If infants have successfully learned the object-label pairings, they are expected to dishabituate during switch trials, with longer looks to the presented object demonstrating that they notice an unexpected mismatch between the label and object. Repeated reproductions of the Switch task have demonstrated that 14-month-olds are apparently able to learn pairs of labels sufficiently to dishabituate during switch trials when those labels differ by multiple phonemes (e.g., "lif" and "neem"), but fail when the labels are lexical neighbors (e.g., "buk" and "puk") (Stager and Werker, 1997; Rost and McMurray, 2009; Yoshida et al., 2009). This result is surprising, considering that infants at this age successfully attend to phonemic contrasts presented with a display of a checkerboard pattern, presumably because this visual information is unlikely to be interpreted as an object which can be labeled (Stager and Werker, 1997; Werker et al., 1998).

Stager and Werker (1997) argued that as infants begin to learn words, the amount of phonetic detail they recruit for speech perception changes. At 14 months, the infants apparently attend to the visual displays, yet demonstrate decreased sensitivity to phonetic detail. Stager and Werker (1997) hypothesized that as infants begin mapping sounds onto meanings, they rely on more abstract, less detailed representations to learn words. However, subsequent modifications to the experimental procedure demonstrate that infants are capable of bringing finer

perception of phonetic detail to bear in this task (Fennell and Werker, 2004; Fennell et al., 2007; Rost and McMurray, 2009; Yoshida et al., 2009; Fennell and Waxman, 2010; Fennell, 2012). This paper focuses on one such experiment showing that exposure to multiple speakers during habituation helps 14-month-old infants succeed on the Switch task.

2.1. Multiple talker variation

Rost and McMurray (2009) demonstrated that exposure to exemplars from multiple voices during habituation increased 14-month old infants' dishabituation in response to misnamings at test. They trained infants on two lexical neighbors ("buk," "puk") in a Switch task, with stimuli recorded either from a single speaker, or from a total of 18 different speakers. Unlike the infants who heard exemplars recorded in a single voice, infants in the condition with multiple speakers successfully discriminated lexical neighbors on the switch trials. Figures 1a and 1b show that the looking time to same trials is significantly less for switch trials where the infants were exposed to multiple talkers.



(a) Single speaker condition **(b)** Multiple speaker condition

Figure 1: Experimental results from Rost and McMurray (2009): 14 month-old infants do not have significantly longer looking time to the labeled object on switch trials after exposure to a single speaker in pre-test (left) but they do after exposure to multiple speakers in pre-test (right).

The authors attributed the infants' success in the multiple-speaker condition to a greater availability of useful phonetic variation in the input. Apfelbaum and McMurray (2011) subsequently modeled these results using associative learning principles. They assumed that cues which tend to have common values across all of a speaker's productions (i.e. cues to the speaker's identity) are attested across both of the presented words. In an associative learning model, when the test token shares attributes with the habituation stimuli, these non-contrastive cues can cause partial activation of both categories and prevent the infant from recognizing the

trial as a misnaming (Apfelbaum and McMurray, 2011). Instead of the infants attending to the acoustic cues which indicate the speech contrast being tested, this account predicts that infants may be attending in their perceptual learning primarily to acoustic cues characteristic of a particular speaker. Attending to speaker-specific cues may cause infants in the single speaker condition to fail to register the phonological contrast between two labels simply because they were spoken by the same person. In the multiple speaker condition, where cues to speaker identity are different with each observed token, it is not possible for the learner to make this mistake.

This explanation for differences between the single speaker and multiple speaker Switch task conditions relies on the supposition that infants are demonstrating immaturity in their knowledge about which phonetic cues are relevant for distinguishing object labels. Rost and McMurray (2009) and Apfelbaum and McMurray (2011) both argued that exposure to a more diverse data set better facilitated the categorical learning by clarifying which cues were relevant for distinguishing the label. However, we show in our simulations that these results can also be explained while assuming that infants have adult-like phonetic perception.

We show that infants' selective inattention to both contrastive and non-contrastive cues can be predicted by incorporating an inference about the reliability of speakers. Specifically, the observation of multiple speakers may license a social inference that agreement among the speakers is indicative of reliability (Fennell and Waxman, 2010). There is substantial evidence that older children perform this type of inference in word learning. In the next section we describe the evidence that reasoning about epistemic trust influences word learning behavior in preschool age children, and then describe a model by Shafto et al. (2012) which captures this process. We then show in the following section that this model can be used to explain the results in Rost and McMurray (2009) as the result of a social inference about the quality of linguistic informants and their testimony.

3. Epistemic trust

Epistemic trust is an account of how learners selectively attend to informants and their testimony. There is a large body of work suggesting that preverbal infants, including newborns, use non-linguistic observations to evaluate social partners (Akhtar and Gernsbacher, 2008; Coulon et al., 2011; Maurer and Werker, 2014; Cirelli et al., 2016), and that infants' representations of informants reflect their previous observations of the informant's reliability (Zmyj et al., 2010; Poulin-Dubois et al., 2011; Tummeltshammer et al., 2014). Judgments of an informant's value as both a linguistic and a non-linguistic source of information are tightly linked in infants (Schachner and Hannon, 2011; Kinzler and Dautel, 2012; Liberman et al., 2017), and an abundance of evidence suggests that processing of indexical and referential information are neurally functionally integrated in adults (see Kuhl, 2011).

Despite its absence from many models of early word learning, the integration of linguistic and non-linguistic perception in infants is well accepted in the literature. The acceptance of this link is expressed with reference to concepts such as perception of an "in-group" which has both characteristic linguistic and characteristic non-linguistic features and behaviors. Studies of the ontogeny of attitudes towards in-group and out-group members (e.g., Buttelmann and Böhm, 2014; Mahajan and Wynn, 2012) are predicated on the assumption that socially motivated cognitive grouping is part of human behavior and likely emerges quite early. The fact that infants use epistemic trust in non-linguistic domains, as well as the evidence for a tight link between linguistic and non-linguistic judgments of informants, suggests that the epistemic trust literature can fruitfully be brought to bear on experiments of early word learning.

Moreover, there is evidence that older children use epistemic trust in ways that, we argue, directly bear on the findings of Rost and McMurray (2009). Corriveau et al. (2009) presented four- and five-year-old children with a display of unfamiliar adults labeling unfamiliar objects. The children were then asked to choose an adult to endorse a label for yet another novel object. When the children observed all but one of the adults agreeing on object labels, they preferred to endorse labels from adults who they had observed participating in the consensus, as opposed to those who they had seen dissenting. Since both the labeled objects and the speakers supplying the labels were unfamiliar to the experimental participants, they could not have been relying on prior knowledge of either when selecting labels to endorse; instead, observing an informant participating in a consensus was enough to influence the children's perception of their testimony.

Shafto et al. (2012) demonstrated that a variant of a model they had previously introduced (Shafto and Goodman, 2008; Shafto et al., 2014), which jointly infers the label for an object and the quality of the informants providing that label, could accurately simulate the word learning experiment from Corriveau et al. (2009). They showed that even without knowing the correct label or the knowledgeability of particular speakers, a listener could reason about how likely a number of informants are to give identical labels, and how likely such agreement is to indicate accuracy in the labeling action. In essence, their explanation proposes that increasing numbers of speakers agreeing on a label drastically decreases children's posterior probability that they are all unknowledgeable. In the next section we demonstrate that this model can also predict the pattern of results in the Switch tasks conducted by Rost and McMurray (2009).

4. Simulating the Switch Task

We use the model from Shafto et al. (2012) of reasoning about categories and speakers of unknown reliability to simulate two experiments from Rost and McMurray (2009), contrasting the behavior of infants habituated to exemplars which were produced by either a single speaker, or by multiple unique speakers.

4.1. Model

In the Switch task experiments we are simulating, infants are habituated to object-label pairings, where the objects are visually distinct and the labels are lexical neighbors ("buk","puk"). They are then tested to see if they successfully differentiate same and switch trials—when the labels are applied to either the same object as seen in test (same trial), or to the object which was assigned its lexical neighbor in habituation (switch trial). Our model simplifies the problem by removing several variables—instead of encoding fine phonetic detail, we assume that the infants are, like adults, capable of categorizing the individual speech tokens. We show that even assuming perfect performance on speech categorization and identifying the referent of the speech act, the source selection model predicts the pattern of results seen in Rost and McMurray (2009).

To model the social inference of our listeners, we begin by supposing they observe some data D, and then infer which is the correct category C. This inference occurs under uncertainty about the reliability of the speaker, K. Figure 2 shows a graphical model for inferring the label and speaker knowledgeability for some number of speakers S who each contribute m utterances.

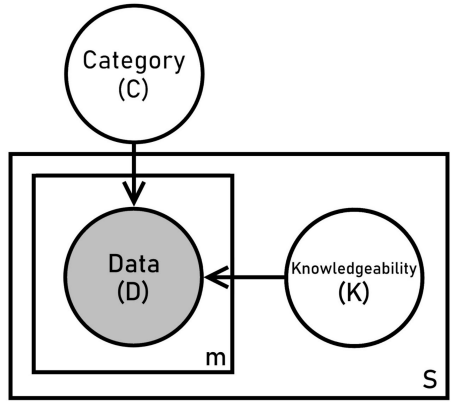


Figure 2: Graphical model of inference about object label C and speaker knowledgeability K given an observed label D

The speaker characteristic of knowledgeability encodes how likely it is that the speaker uses the correct label for the object. We present a simple case, assuming that there are only two levels of knowledgeability: knowledgeable speakers always produce the correct label, whereas only a fraction of unknowledgeable speakers do so. Supposing a uniform prior over these two knowledge states, we set the probability that a given speaker is knowledgeable to ½. This simple binary contrast should be considered compelling, as it illustrates that even with the minimum number of speaker states under consideration, task performance can be predicted by imputing either more or less social differentiation of speakers to the listener.

In effect, the model posits that knowledgeable speakers categorically give trustworthy testimony, and variation in form strictly occurs between unknowledgeable informants. We assume that unknowledgeable speakers will assign an object a label that is uniformly selected from among all of the labels that are under consideration in the word learning task, making the rate of knowledgeable behavior among unknowledgeable speakers $1/n$, where n is equivalent to the number of labels under consideration. Knowledgeable speakers are defined as homogeneously predictable and undifferentiated with respect to labeling behavior. By contrast, unknowledgeable speakers display characteristic patterns of errors which allow most to be differentiated from knowledgeable speakers by their labeling behavior. This means that infants receive a mixture of three kinds of testimony: correct testimony from knowledgeable sources, correct testimony from unknowledgeable sources, and incorrect testimony from unknowledgeable sources.

Given some set of observations, the listener must infer both the correct label and the knowledgeability of each speaker. To model the pair of experiments in Rost and McMurray (2009) we need to do this joint inference under two conditions. In the single speaker condition, all observations are attributed to one source. In the multiple speaker condition, each observation may be attributed to a different source.

4.2. Simulation 1: Single speaker

The participants in Experiment 1 from Rost and McMurray (2009) heard seven consecutive instances of the same exemplar. We simulate the beliefs of the infant at the end of this habituation period by using our model to calculate the joint posterior probability of the category and the speaker's knowledgeability after seven instances of the same unambiguous token. We then compute the extent to which the model believes that the label seen during habituation is the correct label by computing the model's marginal posterior distribution over labels.

We model the infant's looking time as the result of a joint inference on (C_x, K_w) for a sequence of data points $D_{\vec{z}}$. We predict that looking time will correlate with certainty about the speaker's reliability. In this condition, all the data points are associated with a single belief about the knowledgeability of the speaker. Equation 1 gives the joint posterior probability of the category being index x and speaker having knowledgeability w.

$$P(C_x, K_w|D_{\vec{z}}) = \frac{P(D_{\vec{z}}|C_x, K_w)P(C_x, K_w)}{\sum_{x'w'} P(D_{\vec{z}}|C_{x'}, K_{w'})P(C_{x'}, K_{w'})} \quad (1)$$

Although the original experiment employs a two-way contrast, the simulations presented here suppose a three-way phonetic contrast; the qualitative results presented here are not driven by the number of categories under consideration, and we use three for illustrative purposes. The hypothesis space for the single speaker

condition is shown in Table 1. In the next section we illustrate how the size of the hypothesis space grows exponentially with the evaluation of additional speakers.

Table 1: Hypothesis space given a scenario where the object label (C_x) is unknown, and the knowledgeability of one informant (K_1) is unknown.

Possible values for category (C_x) and single speaker knowledgeability (K_1)
C_x=buk, K_1=0
C_x=buk, K_1=1
C_x=puk, K_1=0
C_x=puk, K_1=1
C_x=duk, K_1=0
C_x=duk, K_1=1

4.3. Simulation 2: Multiple speakers

For the condition where the infant hears labels from multiple speakers, we model this as a joint inference on $(C_x, K_{\vec{w}})$ for a sequence of data points $D_{\vec{z}}$. In this condition, each data point is the contribution of a distinct speaker, and as such is associated with a unique belief about knowledgeability specific to this speaker. In other words, for a set of data with m elements, the listener must now infer a sequence $K_{\vec{w}}$ with length m.

$$P(C_x, K_{\vec{w}}|D_{\vec{z}}) = \frac{P(C_x, K_{\vec{w}}) \prod_i P(D_{z_i}|C_x, K_{w_i})}{\sum_{x',w'} P(C_{x'}, K_{\vec{w'}}) \prod_i P(D_{z_i}|C_{x'}, K_{w'_i})} \tag{2}$$

Table 2: Hypothesis space given a scenario where the object label (C) is known, but the knowledgeability of three informants (K_1, K_2, K_3) is unknown.

category and first speaker (C_x, K_1)	second speaker (K_2)	third speaker (K_3)
C_x =buk, $K_1 = 0$	$K_2 = 0$	$K_3 = 0$
		$K_3 = 1$
	$K_2 = 1$	$K_3 = 0$
		$K_3 = 1$
C_x =buk, $K_1 = 1$	$K_2 = 0$	$K_3 = 0$
		$K_3 = 1$
	$K_2 = 1$	$K_3 = 0$
		$K_3 = 1$

The learner must produce one belief about C and beliefs about an indexed set of variables K_1 through K_m. Table 2 illustrates how the size of the hypothesis space grows when observing three speakers. With the evaluation of additional speakers, it would continue to grow exponentially.

This distribution is too complex to calculate analytically, so we instead use Gibbs sampling, a Markov chain Monte Carlo method that allows us to perform approximate inference (Geman and Geman, 1984). Given the conditional distributions, $P(C_x|K_{\vec{w}}, D_{\vec{z}})$ and $P(K_{\vec{w}}|C_x, D_{\vec{z}})$, the Gibbs sampler iteratively samples from these, using the new value obtained at each step to sample from the other conditional distribution. This iterative sampling process converges to approximate the joint distribution described in Equation 2.

4.4. Results

In modeling looking times in the Switch task, we assume that increased certainty about the label, C, is expected to correlate with increased looking to the target image on switch trials, as a result of infants being more surprised at the novel label. By contrast, infants who are unsure of the label should demonstrate lower looking times, because their existing uncertainty makes a new label less surprising. Table 3 shows the posterior probability the model assigned to each label after familiarization, across the two conditions. The table shows the predicted probability of each label, or category, C, being the correct name for the object, given an observation of data, D. In this case, the infant has observed seven tokens of the label "buk" associated with the object during habituation. The model predicts that infants will be much more confident of the object label after hearing multiple speakers agree. This occurs because agreement among speakers increases the model's belief that not all the speakers are unknowledgeable. Since unknowledgeable speakers do not always produce the correct label, this makes it more likely that speakers in agreement are knowledgeable, and have supplied the correct label. Rather than reflecting more sophisticated phonetic perception, improvement on the Switch task at 18 months may reflect infants' developing strategies for determining informant reliability.

Table 3: Posterior probability of label and knowledgeability after seven observations of "buk"

	Single speaker	Multiple speaker
Label	$P(C \mid D)$	$P(C \mid D)$
C="buk"	0.6667	0.9992
C="puk"	0.1667	0.0004
C="duk"	0.1667	0.0004

5. Discussion

In acquiring a specific speech variety, not all sources will be equally useful to a language learner. If the child is rationally interpreting evidence of label variation in a social setting, we should expect their attention to categorical sound variation to be distributed in accordance with their beliefs about the usefulness of speakers. This paper applied a computational model from Shafto et al. (2012) to simulate a pair of Switch task experiments, demonstrating that the pattern of results documented by Rost and McMurray (2009) could be interpreted as evidence of social inference in word learning. Rather than changes in phonetic learning, social development may be the best explanation for improvement on the Switch task. Specifically, we argue that the cognitive development which underpins infant progress on the Switch task may be governed by changes in the meta-cognitive function of epistemic trust.

While the majority of Switch task studies have attempted to measure infant recognition of words by assuming that the representational structures at issue did not reflect any meaningful variation in encoding of details about informants or their reliability, we argue that their results still support the source selection hypothesis. In what follows, we discuss two main alternative explanations that are often given to account for findings in the Switch task literature, and discuss their compatibility with our account. The first is that infants who fail on the Switch task do so due to a resource limitation. The second is that infants who fail on the Switch task do so due to the ambiguity or absence of referential cues. We argue that the source selection hypothesis is a parsimonious explanation which effectively unites both accounts.

5.1. The cognitive load account

Associating an object with a label also requires the coordination of other cognitive processes, including attention, segmentation and inference about the speaker's referential intent. The failure to demonstrate phonemic discrimination on the Switch task has sometimes been attributed to a resource limitation (Stager and Werker, 1997; Pater et al., 2004). The source evaluation account is compatible with these accounts, however it further provides a specific falsifiable prediction regarding the difficulty of the task: word learning is most likely to be evident under task conditions which facilitate inferences that either the informant or their testimony is reliable, excepting in younger infants who may not yet use the same sophisticated epistemic trust processes in their word learning. Developmental changes on the Switch task may be best explained with reference to changing strategies for evaluating informants.

The cognitive load hypothesis is supported by infant improvement in performance on variations on the Switch task which are designed to be easier. Fourteen-month old infants perform above chance on a simpler Switch task administered with a preferential looking paradigm, consistent with the hypothesis that declines in performance observed under other conditions are the result of task difficulties

(Yoshida et al., 2009). This slight preference is predicted by the model given in Apfelbaum and McMurray (2011), and under the assumption that the experimental participant assigns the event that the speaker in the habituation phase is knowledgeable a non-zero probability, our model predicts a preference for the labeled object during a preferential looking test as well.

In tasks involving familiar words and objects, 14-month-olds demonstrate increased sensitivity to phonetic detail (Swingley and Aslin, 2002; Fennell and Werker, 2003, 2004; Fennell, 2012). Supposing at least part of infants' difficulty in succeeding at the Switch task with minimally different labels is attributable to the increased task requirements of the audio-visual associative learning required to respond to novel words, then the presentation of familiar stimuli should alleviate that difficulty. In effect, the participants' prior knowledge facilitates the task.

In our model, we can simulate this contrast by increasing the prior on knowledgeability. The parameter K in our model predicts the likelihood of an informant both correctly identifying and labeling the referent, which, whether familiar or novel, is known to the experimental participant. Assuming that the infant believes that a familiar object is more likely to be known to their interlocutor, or that the object is simply more salient (and thus more likely to be known), an increase in $P(K)$ simulates the effect of familiar stimuli. Rather than the familiarity of the lexical items facilitating lexical processing, it may facilitate epistemic trust in the informant, indirectly resulting in greater phonetic sensitivity.

A systematic comparison of infant performance after exposure to different amounts of testimony from differing numbers of informants would be required to tease apart these effects. The source-tracking hypothesis predicts that infants will prefer the label offered by the majority of pre-test speakers, except when they have observed evidence that those speakers are unreliable.

5.2. The referential ambiguity account

Performance on the Switch task improves when the novel word is embedded in an overtly referential phrase (i.e., "look at the *blick*") (Fennell and Waxman, 2010) or when the training phase contains familiar named objects (Fennell et al., 2007). However, when familiar objects in habituation are paired with exclamations (e.g., "Wow!" or "Whee!"), no improvement is observed (Namy and Waxman, 2000). These results have been interpreted as support for the hypothesis that 14-month-old infants' failure on the Switch task is a consequence of referential ambiguity. Cues which make the stimulus presentation more clearly a referential act increase the likelihood that infants demonstrably create a mapping between the word and object using fine phonetic detail.

Other studies demonstrate that the infants are more likely to succeed on the Switch task when additional referential cues are present. For example, 14-month olds look longer to switch trials after being exposed to pre-test trials showing the speaker labeling familiar objects (Fennell and Waxman, 2010). Fennell and Waxman (2010) interpreted these results to indicate that the infants were assisted in

making inferences about the referential intentions of the speakers. In other words, strong referential cues helped the infants infer that the provided label was intended by the speaker to correspond with the object.

Our hypothesis is consistent with this one, but makes an important distinction in scope. Because Fennell and Waxman (2010) discusses infant inference about the referential nature of observed speech provided by a single individual, these data cannot distinguish between hypotheses which rely on characterizing the infant's perception of the specific informant's intentions and the infant's perception of the label when used by other informants. It is possible that the participants interpreted pre-test trials displaying accurate labeling behavior as evidence that the speaker is not only intentionally engaging in referential acts, but is doing so credibly as a speaker of the infant's linguistic variety. The subsequent improvement in performance for these infants may demonstrate that they formed a belief that beyond their intent to refer, the speaker is epistemically a source of accurate linguistic data. Rather than simply inferring whether the speaker intends to label the object, the participant may also be concerned with whether the label is accurate and likely to be used by other speakers of their language. The latter inference is captured by our model.

The task is designed to make the labeled object salient to the experimental participant, so if we assume, as we have before, that the infants do know which object is being referred to, the inclusion of additional cues that the speech act references this object may again be encoded as an increase in the prior probability on K. Rather than simply tracking the speech acts themselves, a listener who is also sensitive to source may interpret additional referential cues as a reflection on the quality of the linguistic informant. We expect any stimuli which bias the infant to believe the informant is more likely to select both the correct referent and label will also result in them being more surprised on switch trials, consequently improving performance on the task.

5.3. Investigating infant preferences for informants

In order to explore the source evaluation hypothesis, which links processes of epistemic trust to performance on phonetic discrimination tasks, several modifications to the Switch task may be useful in future research.

Firstly, the source-tracking hypothesis predicts that 14-month olds would demonstrate improved phonetic sensitivity on a Switch task featuring novel words recorded by familiar speakers. However, such a result would also be compatible with a cognitive load hypothesis. The task might also be modified to contrast presentation of novel words in carrier phrases recorded by speakers in alternate speech varieties. If foreign-language carrier phrases provide any benefit then this would challenge the source evaluation explanation. However, infant sensitivity to within-language speech variation, e.g. improved or decreased performance on native language carrier phrases presented in different accents, would not necessarily be inconsistent with this hypothesis.

Suppose a habituation featuring labeling from two speakers. The source-tracking hypothesis predicts that whether listeners demonstrate sensitivity to a phonetic contrast will be partially predicted by their belief that the speaker is knowledgeable. Supposing one of the speakers heard in pre-test is more reliable at labeling familiar objects, infants who hear this speaker's voice on test trials should be more likely to attend to switch trials than infants who hear the less reliable speaker's voice at test. Likewise, the use of a pre-test demonstration where the speaker is shown to be more or less reliable using non-linguistic cues (such as indicating with gaze where an object will appear) may diminish the beneficial effect of naming familiar objects pre-test. If infants are attending to the reliability of the speaker, then demonstrations that they are unknowledgeable in other ways may cause the infant to disprefer attending to that informant's phonetic variation.

The source selection hypothesis predicts that attention to phonetic variation may be either increased or decreased with evidence regarding the quality of the source, and depending upon the prior biases and experiences of the listener. For example, Galle et al. (2015) demonstrated that increased acoustic variability assists 14-month-olds in learning phonemically similar words within the Switch task even when the increased variability is provided by a single talker, rather than multiple ones. The variability of stimuli in this experiment were designed to be naturalistic, varying across several speaking styles. The presented acoustic variation is therefore likely similar to the sort infants are accustomed to regularly encountering from reliable linguistic informants in their environment. It is also possible that these results demonstrate a prior bias to consider more highly variable talkers reliable, rather than an asocial sensitivity to acoustic variability. In order to be certain that modifications to the protocol reflect generalizable perceptual tendencies, future Switch task investigations must control for how infants may come to the lab with varying strategies for inferring informant reliability.

Under the source selection account, preferences for informants which are formed early on may have far reaching effects, shaping the development of the lexicon for years afterwards. To explore this hypothesis, it is necessary to conduct a systematic comparison of infant performance after exposure to different amounts of testimony from differing numbers of informants. It is also necessary to determine how allocation of epistemic trust may vary between populations. Children from different cultural backgrounds and learning in different modalities are expected to eventually acquire distinct strategies for determining the reliability of an informant. Therefore, before we may tease apart the effects of exposure and epistemic trust on word learning, we must understand normal variation in its application. The present work suggests a new research program uniting studies of developmental social psychology with psycholinguistic processing, to discover how variation in phonetic representations are affected by the perception of identity, including attributes such as authority, gender and race.

References

Nameera Akhtar and Morton Ann Gernsbacher. On privileging the role of gaze in infant social cognition. *Child Development Perspectives*, 2(2):59–65, 2008.

Keith S. Apfelbaum and Bob McMurray. Using variability to guide dimensional weighting: Associative mechanisms in early word learning. *Cognitive Science*, 35:1105–1138, 2011.

David Buttelmann and Robert Böhm. The ontogeny of the motivation that underlies in-group bias. *Psychological Science*, 25(4):921–927, 2014.

Nick Chater and Christopher D. Manning. Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, 10(7):335–344, 2006.

Laura K. Cirelli, Stephanie J. Wan, and Laurel J. Trainor. Social effects of movement synchrony: increased infant helpfulness only transfers to affiliates of synchronously moving partners. *Infancy*, 21(6):807–821, 2016.

Kathleen Corriveau and Paul L. Harris. Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, 12(3):426–437, 2009.

Kathleen Corriveau, M. Fusaro, and Paul L. Harris. Going with the flow: Preschoolers prefer nondissenters as informants. *Psychological Science*, 20(3):372–377, 2009.

Marion Coulon, Bahia Guellai, and Arlette Streri. Recognition of unfamiliar talking faces at birth. *International Journal of Behavioral Development*, 35(3):282–287, 2011.

Christopher T. Fennell. Object familiarity enhances infants use of phonetic detail in novel words. *Infancy*, 17(3):339–353, 2012.

Christopher T. Fennell and Sandra R. Waxman. What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, 81(5):1376–1383, 2010.

Christopher T. Fennell and J. F. Werker. Infant attention to phonetic detail: Knowledge and familiarity effects. In *Proceedings of the 28th annual Boston University Conference on Language Development*, volume 1, pages 165–176, Somerville, MA, 2004. Cascadilla Press.

Christopher T. Fennell and Janet F. Werker. Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46(2):245–264, 2003.

Christopher T. Fennell, Sandra R. Waxman, and Adriana Weisleder. With referential cues, infants successfully use phonetic detail in word learning. In *Proceedings of the 31st annual Boston University Conference on Language Development*, volume 1, page 206, Somerville, MA, 2007. Cascadilla Press.

Michael C. Frank, Noah D. Goodman, and Joshua B. Tenenbaum. Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, 20 (5):578–585, 2009.

Marcus E. Galle, Keith S. Apfelbaum, and Bob McMurray. The role of single talker acoustic variation in early word learning. *Language Learning and Development*, 11(1):66–79, 2015.

Stuart Geman and Donald Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE-PAMI*, 6:721–741, 1984.

Katherine D. Kinzler and Jocelyn B. Dautel. Children's essentialist reasoning about language and race. *Developmental Science*, 15(1):131–138, 2012.

Katherine D. Kinzler, Kathleen Corriveau, and Paul L. Harris. Children's selective trust in native-accented speakers. *Developmental Science*, 14(1):106–111, 2005.

Melissa A. Koenig and Paul L. Harris. Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, 76(6):1261–1277, 2005.

Patricia K Kuhl. Who's talking? *Science*, 333(6042):529–530, 2011.

Zoe Liberman, Amanda L. Woodward, and Katherine D. Kinzler. Preverbal infants infer third-party social relationships based on language. *Cognitive Science*, 41:622–634, 2017.

Neha Mahajan and Karen Wynn. Origins of "us" versus "them": Prelinguistic infants prefer similar others. *Cognition*, 124(2):227–233, 2012.

Daphne Maurer and Janet F Werker. Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, 56(2):154–178, 2014.

Laura L. Namy and Sandra R. Waxman. Naming and exclaiming: Infants' sensitivity to naming contexts. *Journal of Cognition and Development*, 1(4):405–428, 2000.

Joe Pater, Christine Stager, and Janet Werker. The perceptual acquisition of phonological contrasts. *Language*, 80(3):384–402, 2004.

Steven Pinker. Formal models of language learning. *Cognition*, 7(3):217–283, 1979.

Diane Poulin-Dubois, Ivy Brooker, and Alexandra Polonia. Infants prefer to imitate a reliable person. *Infant Behavior and Development*, 34(2):303–309, 2011.

Gwyneth C Rost and Bob McMurray. Speaker variability augments phonological processing in early word learning. *Developmental Science*, 12(2):339–349, 2009.

Adena Schachner and Erin E Hannon. Infant-directed speech drives social preferences in 5-month-old infants. *Developmental Psychology*, 47(1):19, 2011.

Patrick Shafto and Noah Goodman. Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In *Proceedings of the 30th annual conference of the Cognitive Science Society*, pages 1632–1637, 2008.

Patrick Shafto, Baxter Eaves, Daniel J. Navarro, and Amy Perfors. Epistemic trust: modeling children's reasoning about others' knowledge and intent. *Developmental Science*, 15 (3):436–447, 2012.

Patrick Shafto, Noah D. Goodman, and Thomas L. Griffiths. A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, 71: 55–89, 2014.

Christine L. Stager and Janet F. Werker. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388:381–382, 1997.

Daniel Swingley and Richard N. Aslin. Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13(5):480–484, 2002.

Kristin. S. Tummeltshammer, Rachel Wu, David M. Sobel, and Natasha Z. Kirkham. Infants track the reliability of potential informants. *Psychological Science,*, 25(9):1730–1738, 2014.

Janet F. Werker, Leslie B. Cohen, Valerie L. Lloyd, Marianella Casasola, and Christine L. Stager. Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, 34(6):1289–1309, 1998.

Janet F. Werker, Christopher T. Fennell, Kathleen M. Corcoran, and Christine L. Stager. Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, 3(1):1–30, 2002.

Katherine A. Yoshida, Christopher T. Fennell, Daniel Swingley, and Janet F. Werker. Fourteen month-old infants learn similar sounding words. *Developmental Science*, 12(3): 412–418, 2009.

Norbert Zmyj, David Buttelmann, Malinda Carpenter, and Moritz Daum. The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology*, 106(4):208–220, 2010.

Proceedings of the 44th annual Boston University Conference on Language Development

edited by Megan M. Brown
and Alexandra Kohut

Cascadilla Press Somerville, MA 2020

Copyright information

Proceedings of the 44th annual Boston University Conference on Language Development
© 2020 Cascadilla Press. All rights reserved

Copyright notices are located at the bottom of the first page of each paper.
Reprints for course packs can be authorized by Cascadilla Press.

ISSN 1080-692X
ISBN 978-1-57473-057-9 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com