

Developmental Psychology

Learning to Remember by Learning to Speak

Marc Ettliger, Jennifer Lanter, and Craig K. Van Pay

Online First Publication, June 17, 2013. doi: 10.1037/a0033317

CITATION

Ettliger, M., Lanter, J., & Van Pay, C. K. (2013, June 17). Learning to Remember by Learning to Speak. *Developmental Psychology*. Advance online publication. doi: 10.1037/a0033317

Learning to Remember by Learning to Speak

Marc Ettliger

Veterans Affairs Northern California Health Care System,
Martinez, California

Jennifer Lanter

University of Wisconsin–Green Bay

Craig K. Van Pay
University of Houston

Does the language we speak affect the way we think, and if so, how? Previous researchers have considered this question by exploring the cognitive abilities of speakers of different languages. In the present study, we looked for evidence of linguistic relativity within a language and within participants by looking at memory recall for monolingual children ages 3–5 years old. At this age, children use grammatical markers with variable fluency depending on ease of articulation: Children produce the correct plural more often for vowel-final words (e.g., *shoes*) than plosive-final words (e.g., *socks*) and for plosive-final words more often than sibilant-final words (e.g., *dresses*). We examined whether these phonological principles governing plural production also influence children's recall of the plurality of seen objects. Fifty children were shown pictures of familiar objects presented as either singular or multiple instances. After a break, they were required to indicate whether they saw the singular- or multiple-instance version of each picture. Results show that children's memory for object plurality does depend on the phonology of the word. Subsequent tests of each child's production ability showed a correlation between a child's memory and his or her ability to articulate novel plurals with the same phonological properties. That is, what children can say impacts what they can remember.

Keywords: language, memory, child language acquisition, phonology, plural

Supplemental materials: <http://dx.doi.org/10.1037/a0033317.supp>

Does the language we speak shape the way we think? This question, posed most explicitly by Benjamin Lee Whorf and Edward Sapir (Whorf, 1956), has motivated decades of research exploring the ways in which memory, attention, and perception differ among speakers of different languages (Boroditsky, 2001; Fausey & Boroditsky, 2011; Frank, Everett, Fedorenko, & Gibson, 2008; Gentner & Goldin-Meadow, 2003; Roberson, Davies, & Davidoff, 2000). The properties of language not only vary from language to language, however; they also vary developmentally, with children showing different levels of linguistic competence from birth well into adolescence (Brown, 1973). One well-documented example is the protracted development for the use of the English plural (Berko, 1958; Ettliger & Zapf, 2011). In the

present study, we explored how children's memory for whether they saw single or multiple instances of an object depends on their language ability. Furthermore, whereas the Sapir–Whorf hypothesis, or *linguistic relativity*, has traditionally been used to denote the influence of semantic categories on thought (e.g., Bowerman & Levinson, 2001), there is a wide range of linguistic knowledge and linguistic abilities that may also influence cognition and perception. In the present study, then, we considered a novel linguistic property, phonology, and its influence on thought. In particular, we asked whether children's plural memory depends on fluency with plural production.

Research on linguistic relativity has generally focused on cross-linguistic differences in categorization and perception (e.g., Bowerman & Levinson, 2001), but there is also strong evidence that the ability to remember can also depend on properties of the native language of the participant. For example, Fausey and Boroditsky (2011) compared Spanish and English speakers' recollection of who did an action based on the observation that English speakers tend to use agentive language for accidental events (e.g., *She broke the vase*), whereas Spanish speakers tend to use the passive construction (e.g., *The vase broke*). English speakers recalled the agent more often than Spanish speakers, and similar effects are also found for cross-linguistic differences in memory for colors (Roberson et al., 2000) and motion scenes (Slobin, 2003).

There is also evidence that the presence or absence of number terms (e.g., *two*, *three*, and so on) in a language can impact cognitive processing (Dehaene, Spelke, Pinel, Stanescu, &

Marc Ettliger, Research Service, Veterans Affairs Northern California Health Care System, Martinez, California; Jennifer Lanter, Department of Human Development and Psychology, University of Wisconsin–Green Bay; Craig K. Van Pay, Department of Psychology, University of Houston.

This work is supported by National Institutes of Health Grant T32 NS047987 to Marc Ettliger. We wish to thank Linda B. Smith, Carla Hudson Kam, David Barner, and Amy Finn for comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Marc Ettliger, Research Service (151/MTZ), Veterans Affairs, Northern California Health Care System, 150 Muir Road, Martinez, CA 94553. E-mail: ettlinger@gmail.com

Tsivkin, 1999). Speakers of the language spoken by the Pirahã lack the ability to express exact quantity (Frank et al., 2008; Gordon, 2004). Although a strong form of linguistic relativity would predict that the Pirahã would be limited in their ability to evaluate exact quantities, Pirahã speakers have been shown to have the ability to match large quantities of objects based on number (Frank et al., 2008). However, the lack of linguistic number does negatively impact their ability to make matches that require the use of memory. Frank et al. (2008) argued that linguistic number serves as a cognitive technology, with language acting as a tool for creating mental representations that facilitate subsequent memory processing. Although recent studies have challenged the notion that the Pirahã can compare quantities with the same accuracy as speakers of numeric languages (C. Everett & Madora, 2012), the notion that language may function as a cognitive tool, either for memory or for comparison purposes, is still a useful hypothesis.

Findings of researchers who have argued against the strong form of linguistic relativity (e.g., Li, Abarbanell, Gleitman, & Papafragou, 2011; Li & Gleitman, 2002; Papafragou, Hulbert, & Trueswell, 2008; Papafragou, Massey, & Gleitman, 2002) are generally compatible with the language as cognitive technology hypothesis. Such studies generally have shown that lacking a certain linguistic term leads to poorer performance in cognitive tasks that may require use of that term but that participants are still able to conceptualize or make use of the concept the term represents. For example, Li et al. (2011) showed that speakers of Tenejapan Mayan, who lack egocentric direction terms (e.g., *left*, *right*), do perform worse on orientation tasks that are egocentric but still perform above chance.

There is, indeed, a vast body of literature exploring and reviewing the relationship between language and thought and language and memory, which is beyond the scope of the present article (see, for example: Gleitman & Papafragou, 2005; Regier & Kay, 2009; Wolff & Holmes, 2011). Notably, much of this research has focused on differences between speakers of different languages (e.g., in the studies cited, Spanish vs. English or Pirahã vs. English). However, people also vary in their linguistic competence over the course of language development, from birth to adulthood, with children learning new grammatical structures well into adolescence. There is ambiguity in how these developmental changes impact mental representations however. For example, the neural representation of color categories in infants changes as they learn color words (Franklin et al., 2008), but it is unclear the degree to which this impacts behavior as Franklin, Clifford, Williamson, and Davies (2005) argued that there is no impact of color terms on memory. In fact, even as adults, the relationship between language and color cognition is complex, with the magnitude of language influence differing between left and right visual fields, likely due to the lateralization of the lexicon (Gilbert, Regier, Kay, & Ivry, 2006). Thus, there not only appears to be differences in the relationship between language and color categorization across languages and across developmental periods, but even the eye used in perception can be a confounding factor.

The relationship among language, thought, and memory has also been examined through analysis of the morphological competence in children. Specifically, Li, Ogura, Barner, Yang, and Carey (2009) suggested that acquiring plural morphology is not, in fact, necessary for learning the singular–plural distinction. More generally, studies in which the influence of language on memory in

children has been specifically examined suggest little or no effect before the age of 6 years (Gathercole, 1998), as children are argued to not subvocalize, for example, and thus do not substantively differ in their ability to remember long versus short words (Yuzawa, 2001).

Experiments comparing adults with children or even children of the same age with different grammatical abilities would not reveal the causal relationship between language and thought because of the parallel developmental trajectory of language and other cognitive abilities. That is, adults generally have better linguistic competence than children and also have better memory, reasoning, and attentional abilities. We cannot draw conclusions about the influence of language on thought given these positive correlations, however.

This same issue of indeterminate causation affects cross-linguistic research as well. For example, in the case of the agent recall in the experiment contrasting English and Spanish speakers (described earlier), it may be that English speakers care more about agency than Spanish speakers because of culture and that may have led to both the preference for the active grammatical construction in the English language and better recall. The third factor of culture is also proposed as the mechanism explaining the relationship between language and thought among the Pirahã, with culture argued to influence both (D. L. Everett, 2005). Thus, cross-linguistic studies often cannot make claims about whether language is influencing thought, thought is influencing language, or some third factor, such as culture, is influencing both.

We can address this issue by recognizing that grammatical competence in children depends not only on their learning some aspect of grammar but also on their ability to use or say the grammatical marker. This is a novel way of exploring the influence of language on thought as previous studies, including those previously listed, generally focused on presence and absence or the meaning of certain semantic categories. By looking at phonology and articulatory ability, we introduce the notion that nonsemantic domains may affect cognition as well.

A prime example of this idea is children's use of the English plural as mediated by phonology and ease of articulation. Children begin to successfully use the *-s* plural (*key–keys*) before the *-es* plural (*watch–watches*; Berko, 1958), and they use the *-s* plural on vowel-final words (*key–keys*) before consonant-final words (*clock–clocks*; Ettliger & Zapf, 2011). The reason for the *keys–clocks* difference is the challenge associated with saying (MacNeilage & Davis, 2000) or hearing complex codas—a syllable ending in multiple consonants—compared with simple codas—a syllable ending in a single consonant. The plural for sibilant-final words is even more complex, requiring schwa-insertion and an additional syllable (*-es*), as well as being a far less frequent variant of the English plural used in the most restricted context (Berko, 1958).

We considered whether these differences in children's plural production impact their ability to think about multiple objects by testing their memory of objects previously seen. Our goal was to assess the hypothesis that what is part of an individual child's language impacts his or her memory. By looking within subjects, we controlled for culture, development, and the other confounds that are part of the cross-linguistic research that we have discussed. Based on previous research on child speech, we expected that children would be more likely to remember the plurality of vowel-

final words, followed by plosive-final (i.e., nonsibilant consonants such as *p*, *t*, *k*), and then sibilant-final words. In other words, we predicted children's ability to recall the plurality of different objects would be based on the phonology of the word, with words that are articulatorily more difficult to pronounce in the plural being harder to remember.

We expanded this analysis by considering individual differences as well, looking at the production and memory abilities of each individual child for each type of consonant. To do so, we also considered the correlation between children's ability to recall the plurality of sibilant-final words and their ability to articulate the plural for sibilant-final words, their recall and articulation of plosive-final words, and their recall and articulation of vowel-final words. If these correlations are significant, it would support the hypothesis that what children can remember and, thus think, depends, in part, on what they can say.

Method

Participants

Participants were 50 children (27 girls, 23 boys) between 30 and 56 months old ($M = 39.5$ months, $SD = 4.3$) from monolingual English-speaking families.

Materials

A list of 36 imageable nouns was created (see Table 1) for the memory test. Twelve words were vowel-final, 12 sibilant-final (*s*, *sh*, *j*, *ch*) and 12 voiceless-plosive-final (*p*, *t*, *k*). All words were regular plurals known by over 90% of 30-month olds according the MacArthur Communicative Developmental Inventory (MCDI) norms for receptive vocabulary (Fenson et al., 1994). Eight of 12 words for each condition were monosyllabic and four disyllabic. These words were selected to increase the number of words available, as there were an inadequate number of monosyllabic words in each condition matching the described criteria.

For the subsequent production test, a set of 18 nonce words, with the same phonological properties as the memory-test words, were created (Table 2). Each word was paired with a picture of a novel, unfamiliar object and was presented in random order.

Procedure

All participants were tested in a university lab or at a local day care facility and received a children's book for participating in the experiment.

The memory test began with a training/familiarization phase where participants saw pictures of 36 objects as either single or multiple (four) instances (see Figure 1a). Half of the pictures were of a single instance and half were of a multiple instance set, balanced for phonological form and presented individually in random order. The experimenter pointed to each picture and said, "Look here!" to direct participants' attention. The picture was shown for 3 s beginning from when the participants first looked at the picture. Participants were not asked to say anything; in particular, the children were not prompted to say the names of the objects. After training, the children listened to classical music for 4 min and colored to increase task difficulty.

Table 1
Words Used in Memory Experiment Grouped by Ending With d' Memory Performance and Singular and Plural Corpus Frequency From the Brent Corpus of Child-Directed Speech, Ratio and Log Ratio

Condition/word ending	d'	Frequency			Ratio of logs
		Singular	Plural	Ratio	
Vowel-final					
Bee	1.35	84	35	0.42	0.80
Boy	0.55	1,284	28	0.02	0.47
Cow	1.79	243	39	0.16	0.67
Eye	1.63	173	145	0.84	0.97
Key	0.60	13	73	5.62	1.67
Shoe	0.76	183	328	1.79	1.11
Toe	1.96	115	159	1.38	1.07
Tree	2.01	168	79	0.47	0.85
Cookie	1.52	161	47	0.29	0.76
Monkey	1.14	49	10	0.20	0.59
Pizza	1.30	2	0	0.00	0.00
Window	1.40	141	13	0.09	0.52
Plosive-final					
Boat	0.34	54	7	0.13	0.49
Book	1.05	996	189	0.19	0.76
Boot	1.08	3	7	2.33	1.77
Coat	0.55	48	1	0.02	0.00
Cup	0.04	234	11	0.05	0.44
Duck	1.47	172	32	0.19	0.67
Hat	1.67	202	3	0.01	0.21
Sock	1.53	121	147	1.21	1.04
Basket	1.11	112	9	0.08	0.47
Bucket	0.55	44	11	0.25	0.63
Carrot	2.02	23	49	2.13	1.24
Firetruck	1.35	5	1	0.20	0.00
Sibilant-final					
Brush	0.89	42	0	0.00	0.00
Bus	1.15	203	0	0.00	0.00
Dress	0.58	12	1	0.08	0.00
Horse	1.47	87	10	0.11	0.52
Hose	0.86	11	0	0.00	0.00
Juice	1.06	453	2	0.00	0.11
Nose	0.35	554	2	0.00	0.11
Purse	0.89	25	0	0.00	0.00
Necklace	0.52	3	0	0.00	0.00
Orange	0.50	150	19	0.13	0.59
Sandwich	1.39	13	0	0.00	0.00
Toothbrush	0.54	19	1	0.05	0.00

To avoid attention difficulties, each child participated in only 18 trials during the testing phase. A two-alternative forced-choice task was used to display the singular and multiple instance pictures of each object from the training/familiarization phase on a specially designed board (see Figure 1b). Participants were prompted with the question, "Which picture did you see?" and were asked to place one of the two pictures in the center of the board. The trial lasted until the participants placed one picture in the box. Neutral feedback was given after each trial.

Prior to exposure and the 18 test trials, participants were given practice consisting of exposure to two pictures and two practice test trials with right/wrong feedback given by reshewing the training page.

A production task was used to test their ability to produce the plural. Each participant was given 18 trials of a *wug* test (Berko,

Table 2
Nonce Words Used for Plural Production Wug Test Grouped by Ending

Vowel-final ending	Plosive-final ending	Sibilant-final ending
ree	wug	dus
kwee	zeg	tass
koe	lib	nush
vue	geed	niz
metoe	nupick	fipdis
foony	blicket	keedouch

1958). Each nonce word (Table 2) was paired with an unfamiliar photograph (see Figure 2). Participants were presented with a picture of a singular instance of the novel item (see Figure 2a), and the experimenter said the name of the object. The participant was then presented with a picture of two of the objects (see Figure 2b), and the experimenter prompted the participant by telling the child, "Look, now what's here? What do you see?" and encouraging the child to complete the sentence. The production test began with one practice trial with a familiar object where feedback was given.

All testing sessions were video recorded for scoring purposes. Scoring was done by an independent experimenter, and 25% of the participants' responses for both the memory and production tests were scored by an additional experimenter with more than 99% agreement. Memory scoring was based on which picture was first placed in the center box. Production scoring was based on whether the plural morpheme was at all audible in the participants' productions.

Parents of the participants were asked to fill out the MCDI to record the participants' productive vocabulary ($M = 588$; $SD = 83$; range = 439–691).

Results

To assess recall, we used the d' measure of sensitivity from signal detection theory for each participant for each condition. The d' measure combines hits (child selected plural when child saw plural) with correct rejections (child selected singular when child saw singular) instead of simply using percentage correct on plural items, wherein a d' of 0 indicates guessing or no recall of the plurality of any images. This serves to mitigate possible biases for singular or plural responses and is a more effective way of determining children's sensitivity to a plural stimulus, given the large

amount of noise characteristic of children's performance in behavioral studies (Macmillan & Creelman, 2005). Furthermore, 13 children did respond with either all singular or all plural responses to all of the items, indicating they did not understand the task. These participants were excluded from the present analyses. Results including these 13 participants and using percentage correct instead of d' are shown in the supplemental online materials. An analysis of percentage correct reflects the same pattern of results.

In all conditions, children performed well overall, remembering the plurality of the words significantly above chance—vowel-final: $t(36) = 8.9, p < .001$; plosive-final: $t(36) = 6.1, p < .001$; and sibilant-final: $t(36) = 3.8, p < .001$. A one-way within-subject analysis of variance with phonological form as a fixed factor (vowel-final, plosive-final, sibilant-final) and word and participant as random factors were used to analyze d' performance on the memory test. As shown in Figure 3, there is a significant effect of the phonological ending of the word, $F(2, 72) = 8.6, p < .001$. Paired post hoc t tests show that the plurality of vowel-final words (e.g., *shoes*; $d' = 2.3$, standard error [SE] = 0.27) is remembered more often than plosive-final words (e.g., *socks*; $d' = 1.8, SE = 0.29$), $t(36) = 2.2, p = .036, d = 0.73$; the plurality of plosive-final words is remembered more often than sibilant-final words (e.g., *dresses*; $d' = 1.2, SE = 0.32$), $t(36) = 2.1, p = .045, d = 0.70$; and the plurality of vowel-final words is remembered more often than sibilant-final words, $t(36) = 3.9, p < .001, d = 1.3$.

Overall memory performance was moderately, but not significantly, correlated with age, $r(36) = .23, p = .16$, and vocabulary, $r(36) = .31, p = .060$, as measured by the MCDI (Fenson et al., 1994). Breaking down performance by phonological condition shows no significant correlations with age—vowel-final: $r(36) = .16, p = .33$; plosive-final, $r(36) = .19, p = .25$; and sibilant-final: $r(36) = .28, p = .09$. There is, however, a telling difference in correlations between memory performance and vocabulary when the data are broken down by phonological form (see Figure 4). Specifically, there is a significant correlation for sibilant-final words, $r(36) = .44, p = .007$, and for plosive-final words, $r(36) = .36, p = .029$, and no correlation with vowel-final words, $r(36) = -.09, p = .59$. This suggests that memory mirrors the development of plural production, where children first develop mastery of the pluralization of vowel-final words but still struggle with sibilant-final words, with plosives somewhere in the middle.

We also examined the role word frequency and the ratio of singular to plural frequency (extracted from the Brent corpus of

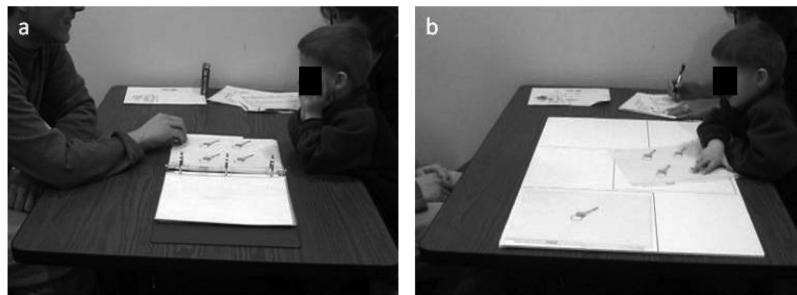


Figure 1. (a) The experimenter shows the child 36 pictures—half single-instance sets, half multiple-(four)-instance sets—in a book, and (b) the child is later tested on 18 of the photos seen earlier by moving the picture he or she saw into the middle, lower box.

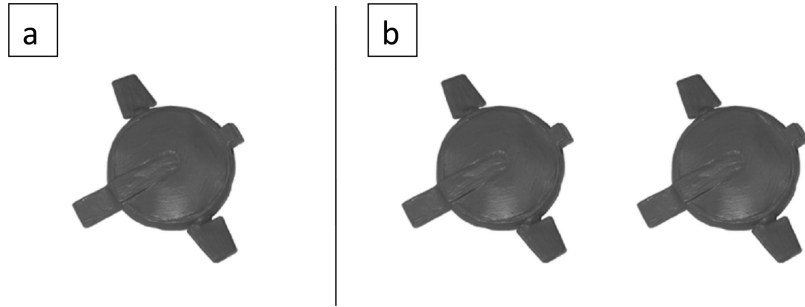


Figure 2. Sample stimuli used for the production test. The singular-denoting page (a) was presented along with one of the words from Table 2. The child was then asked to say the appropriate word for the picture in the plural-denoting page (b).

child-directed speech (Brent & Siskind, 2001; see Table 1) has on children’s memory performance. Indeed, the ratio of the log of singular and plural frequencies are significantly different between conditions, $F(2, 33) = 9.6, p < .001$, suggesting frequency may play a role, though the ratio of raw frequencies is not significantly different, $F(2, 33) = 2.3, p = .11$. There is, however, no evidence for a difference between memory for words that are more frequent in the plural than in the singular compared with the words that are more frequent in the singular than the plural, unpaired t test: $t(6.4) = 1.1, p = .31, d = 0.37$.

We also conducted an analysis of covariance of memory performance with plural:singular and $\log(\text{plural}):\log(\text{singular})$ ratios as continuous variables and word ending as a categorical factor to assess whether word ending still significantly impacts memory after frequency is controlled. The results show that ratio, $F(1, 31) = 0.14, p = .70$, and log ratio, $F(1, 31) = 0.43, p = .52$ are not significant, but word ending remains marginally significant, $F(2, 32) = 3.1, p = .058$. This finding suggests there may be some impact of frequency, but word ending nevertheless has a marginally significant effect beyond that. We discuss this point further in the Discussion section.

Finally, we also considered the role each child’s individual ability to produce the different plural forms relates to his or her memory ability. After the memory test, each child participated in

a *wug* test (Berko, 1958) assessing their ability to produce the plural for nonce words, with a focus on comparing vowel-final, plosive-final, and sibilant-final words. Juxtaposing performance on the memory test with each child’s production ability more directly assesses whether the ability to produce certain grammatical markers impacts cognitive abilities. In line with previous research, children produced the correct plural form less often for sibilant-final words (35%) than for the plosive- (66%), $t(36) = 3.3, p = .002, d = 1.1$, and vowel-final (82%), $t(36) = 5.7, p < .001, d = 1.9$, words and produced the correct plural form more often for vowel-final than for plosive-final words with marginal significance, $t(36) = 2.0, p = .056, d = 0.67$. It is important to note that there is a correlation between a child’s performance in producing the plural and his or her ability to recall plurality for sibilant-final words, $r(36) = .37, p = .02$, plosive-final words, $r(36) = .42, p = .008$, and vowel-final words, $r(36) = .39, p = .014$ (Figure 5). The latter two are significant even after correcting for multiple comparisons ($p < .017$), whereas no other correlations are significant (e.g., vowel production and sibilant production).

To summarize, there is a significant difference in children’s ability to recall the plurality of different objects based on the phonology of the word, with words that are articulatorily more difficult to pronounce in the plural harder to remember. Furthermore, for each individual child, there is a correlation between his or her ability to

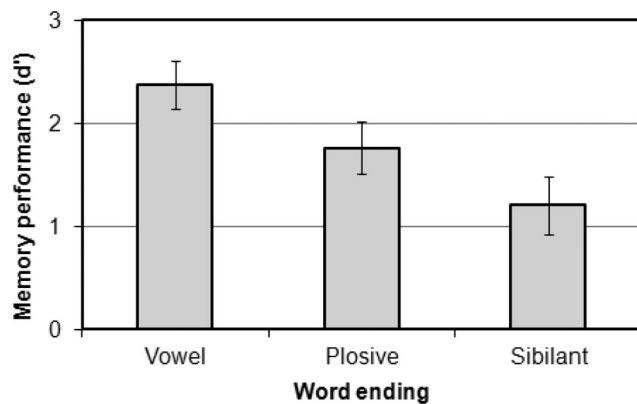


Figure 3. Graph showing d' measure of performance on recall of object plurality based on the phonological form of tested words. Error bars indicate standard error.

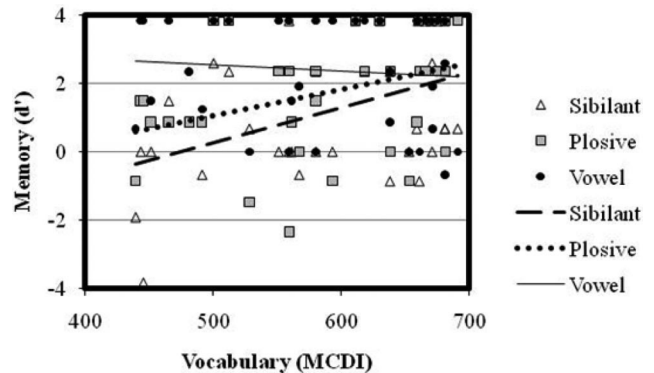


Figure 4. Graph showing d' measure of performance on memory test broken down by condition compared with a MacArthur Communicative Developmental Inventory (MCDI) measure of vocabulary for each child.

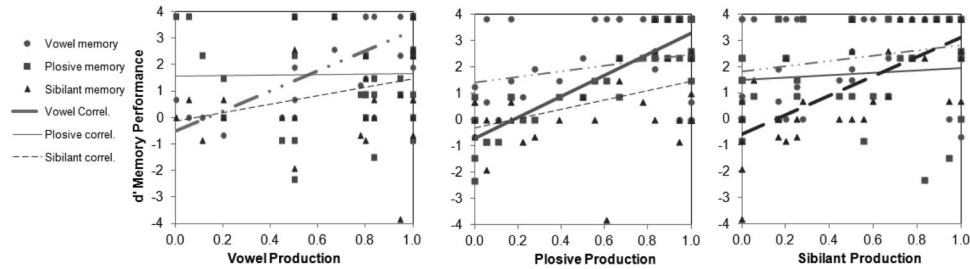


Figure 5. Recall in each condition compared with production ability for each condition. Correlation for each memory condition is strongest with the corresponding production condition, indicated by the bolded line.

produce the plural for words with a certain ending and his or her ability to recall the plurality of objects with that ending.

Discussion

Previous research has shown a relationship between language and thought wherein the presence or absence of some grammatical marker or category impacts some general cognitive ability, such as memory. In the present study, we showed this same effect for children with respect to recalling the number of seen objects based on the phonology of the plural words. The difference in performance across the three conditions—sibilant-final words (e.g., *dresses*), plosive-final words (e.g., *socks*), and vowel-final words (e.g., *shoes*)—shows that children are better able to recall the plurality of objects they saw for words where the plural is easier to pronounce. This effect holds at the level of the individual as well, wherein the ability to recall the plurality of words from each condition correlates with a child's ability to say the plural of nonce words with the same phonological properties. This finding suggests that not only is there a difference in word recall based on the word's articulatory difficulty but that there is also a relationship in individuals, where a child's ability to articulate certain sounds relates to his or her ability to remember them.

Thus, we believe this is the first study to show an effect of language on memory in young children. In previous research, it has been argued that working memory in young children is language independent, with subvocalization strategies and increased verbal fluency beginning around the age of 6 years (Gathercole, 1998; Yuzawa, 2001). While some studies have shown possible indirect linkages between language and memory beginning as early as 4 years old (Gathercole, Pickering, Ambridge, & Wearing, 2004), our study is the first to find direct evidence for this link.

We believe this is also the first study to consider the impact of phonology on cognition. Linguistic relativity is generally associated with the influence of semantic, and at times syntactic, properties on cognition and perception. By looking at production fluency, we therefore explored a novel way in which language may influence thought.

The current results provide insight into linguistic relativity unavailable from previous studies that analyzed data across languages (Boroditsky, 2001; Frank et al., 2008; Gentner & Goldin-Meadow, 2003; Gleitman & Papafragou, 2005; Papafragou et al., 2008). By comparing individuals who speak a single, common language, a number of candidate explanations for the interaction between language and thought were eliminated: the difference in memory for one

set of words versus the other cannot be ascribed to cultural or cognitive differences between different populations or to testing environment. Instead, the difference can only be due to the words themselves and, more likely, the phonology of the words. With respect to causality, given the conditions in this study were phonological and therefore linguistic in nature, we have clear evidence that language is influencing thought. That is, unlike the previous examples of differences in agent recall between English and Spanish speakers (Fausey & Boroditsky, 2011) and differences in numerical recall between Pirahã and English speakers (Frank et al., 2008), wherein culture may be influencing both language and thought or thought may be influencing language, the differences presented here represent a clearer case of language influencing thought.

One alternate explanation we considered is the role that word frequency has on children's memory performance, given that frequency has been shown to have an effect on adults' speed of processing of the plural form across a number of different languages (New, Brysbaert, Segui, Ferrand, & Rastle, 2004). While there is no evidence that frequency does play a role, it does seem to account for some of the effects we found between conditions. In particular, sibilant-final words (e.g., *dresses*) occur relatively less frequently in the plural compared with vowel- (e.g., *shoes*) and plosive-final words (e.g., ratio of *dresses:dress* is .08 while the ratio of *shoes:shoe* is 1.8). It is crucial to note, however, that frequency and production difficulty are not independent measures as it has been oft noted that things that are easier to say occur more frequently in language (Langfitt et al., 1986; Zipf, 1949). Thus, while we would ideally prefer to disentangle the effects of phonology and frequency, the two are inextricably linked. In fact, frequency may serve as one of the mechanisms by which language influences memory as it is quite striking that of the most common words known by children, words for objects that come in groups, are often vowel-final (e.g., *toes, keys, eyes*). Future research can explore this question further.

Frequency may also play a role at a broader level as the *-s* suffix occurs more frequently than the *-es* suffix, possibly accounting for poorer recall of sibilant-final words, which all take the *-es* suffix. However, while this could account for the difference between sibilant- and plosive-final words, it does not account for the difference between vowel- and plosive-final words. More generally, even if word frequency could explain the difference in memory performance, it would only obviate the articulatory explanation we highlighted in favor of a frequency-based one. The results would still implicate language—either the frequency of different allomorphs or the frequency of certain words in the singular or

plural form—as affecting memory. Thus, while there is evidence that some of the observed memory differences may be due to frequency, this still represents evidence that memory for plurality is driven by linguistic factors.

Future research can further address the distinction between strong forms of linguistic relativity, where language determines what can be conceptualized, and weaker forms of the hypothesis, such as “thinking for speaking,” where speakers must attend to certain aspects of the world to speak their language (Slobin, 2003), and even more modest claims where linguistic constraints simply act as minor and ephemeral adjustments on cognitive processing (Gleitman & Papafragou, 2005). The present study does suggest that language is acting as an augmenter to cognition, facilitating recall (Fenson et al., 1994; Gentner & Goldin-Meadow, 2003), rather than limiting thought. Those who subscribe to the strong form of linguistic relativity with respect to number and phonology would argue that children cannot even think about the plurality of sibilant-final words because they lack the articulatory ability; this is unlikely in light of evidence that Mandarin- and Japanese-speaking children can make distinctions based on number before acquiring plural marking (Li et al., 2009). Other studies of plural recall and identification for children ages 2–5 years (Zapf & Smith, 2008, 2009) do show children successfully distinguishing single- and multiple-instance sets. Given these findings, we may dismiss the observation in this study that performance is not above chance for many children ($d' \sim 0$ in Figure 5 for children with production performance below $\sim 30\%$); instead, poor performance is likely due to task difficulty rather than to children not distinguishing single and multiple instances altogether.

Thus, these findings are compatible with the notion of cognitive fluency, such that the ease with which information can be processed impacts a range of diverse cognitive processes including judgments of truth, calculations of frequency, memory, and preference (Alter & Oppenheimer, 2006; Hillary, Moelter, Schatz, & Chute, 2001; Oppenheimer, 2008). For example, stocks with names that are easier to pronounce perform better (Alter & Oppenheimer, 2006), and recall of information is impacted by how easy it is to read the font in which it is presented (Oppenheimer, 2008). In the present study, we find evidence that things that are easier for children to say are also more easily remembered. Future research, which can include adults and a wider range of words with greater variation in difficulty of pronunciation, may provide insight into the specific parameters of this relationship between memory and ease of articulation. For example, it may be that at extremes, words that are very hard to pronounce may be more easily remembered because they stand out more (Zipf, 1949).

Finally, another crucial question concerns the mechanism of how language influences thought and, in particular, how phonology influences memory in children. One possibility is that it involves subvocalization of the stimulus by the children, the idea being that children that can better articulate a subvocalized stimulus and, thus, can more accurately encode its memory. Indeed, this is a hypothesized mechanism for phonological effects on verbal working memory wherein adult participants who are prohibited from subvocalizing, by virtue of having to articulate irrelevant information, perform worse on a verbal working-memory task (Baddeley & Hitch, 1974; Cole & Young, 1975).

There is reason to believe this is not the case in the present study, however. First, the task used here is a visual memory task,

and the subvocalization effect effectively disappears for visual working memory (Salamé & Baddeley, 1982). Furthermore, it has been suggested that children do not subvocalize (Gathercole, 1998; Yuzawa, 2001). Experiments wherein participants are forced to vocalize the stimulus they see, or are inhibited from vocalizing by being forced to vocalize an irrelevant stimulus, can test this hypothesis directly. If subvocalization does not appear to be a factor, there are a number of other hypotheses and questions to consider. Is the observed effect associated more with the encoding of memory, the maintenance of memory storage, or with memory recall? In all cases, does improved articulatory ability sharpen the mental representation of associated words? Does the effect persist over longer periods of time, where working memory, which is modality specific (Baddeley, 1992; Baddeley, Thomson, & Buchanan, 1975), is less relevant and long-term memory plays a great role? While this study was not designed to address these questions, future studies can hopefully provide insight into the mechanisms by which child language development impacts memory.

References

- Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *PNAS: Proceedings of the National Academy of Sciences of the United States of America*, *103*, 9369–9372. doi:060107110310.1073/pnas.0601071103
- Baddeley, A. D. (1992, January 31). Working memory. *Science*, *255*, 556–559. doi:10.1126/science.1736359
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York, NY: Academic Press.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*, 575–589. doi:10.1016/S0022-5371(75)80045-4
- Berko, J. (1958). The child’s learning of English morphology. *Word*, *14*, 150–177.
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers’ conceptions of time. *Cognitive Psychology*, *43*, 1–22. doi:10.1006/cogp.2001.0748
- Bowerman, M., & Levinson, S. C. (Eds.). (2001). *Language acquisition and conceptual development*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511620669
- Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*, 33–44.
- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Cole, R. A., & Young, M. (1975). Effect of subvocalization on memory for speech sounds. *Journal of Experimental Psychology: Human Learning and Memory*, *1*, 772–779. doi:10.1037/0278-7393.1.6.772
- Dehaene, S., Spelke, E., Pineda, P., Stanescu, R., & Tsivkin, S. (1999, May 7). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*, 970–974. doi:10.1126/science.284.5416.970
- Ettlinger, M., & Zapf, J. A. (2011). The role of phonology in children’s acquisition of the plural. *Language Acquisition*, *18*, 294–313.
- Everett, C., & Madora, K. (2012). Quantity recognition among speakers of an amumeric language. *Cognitive Science*, *36*, 130–141. doi:10.1111/j.1551-6709.2011.01209.x
- Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology*, *46*, 621–646. doi:10.1086/431525
- Fausey, C. M., & Boroditsky, L. (2011). Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic Bulletin & Review*, *18*, 150–157. doi:10.3758/s13423-010-0021-5

- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (with Tomasello, M., Mervis, C. B., & Stiles, J.) (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, Serial 242, 59*, No. 5. doi:10.2307/1166093
- Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition, 108*, 819–824. doi:S0010-0277(08)00104-210.1016/j.cognition.2008.04.007
- Franklin, A., Clifford, A., Williamson, E., & Davies, I. (2005). Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology, 90*, 114–141. doi:10.1016/j.jecp.2004.10.001
- Franklin, A., Drivonikou, G. V., Clifford, A., Kay, P., Regier, T., & Davies, I. R. (2008). Lateralization of categorical perception of color changes with color term acquisition. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 105*, 18221–18225. doi:10.1073/pnas.0809952105
- Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, 39*, 3–27. doi:10.1017/S0021963097001753
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*, 177–190. doi:10.1037/0012-1649.40.2.177
- Gentner, D., & Goldin-Meadow, S. (2003). *Language in mind; advances in the study of language and thought*. Cambridge, MA: MIT Press.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 103*, 489–494. doi:10.1073/pnas.0509868103
- Gleitman, L. R., & Papafragou, A. (2005). Language and thought. In K. Holyoak & R. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 663–661). New York, NY: Cambridge University Press.
- Gordon, P. (2004, October 15). Numerical cognition without words: Evidence from Amazonia. *Science, 306*, 496–499. doi:10.1126/science.1094492
- Hillary, F. G., Moelter, S. T., Schatz, P., & Chute, D. L. (2001). Seatbelts contribute to location of lesion in moderate to severe closed head trauma. *Archives of Clinical Neuropsychology, 16*, 171–181.
- Langfitt, T. W., Obrist, W. D., Alavi, A., Grossman, R. I., Zimmerman, R., Jaggi, J., . . . Patton, D. R. (1986). Computerized tomography, magnetic resonance imaging, and positron emission tomography in the study of brain trauma: Preliminary observations. *Journal of Neurosurgery, 64*, 760–767. doi:10.3171/jns.1986.64.5.0760
- Li, P., Abarbanell, L., Gleitman, L. R., & Papafragou, A. (2011). Spatial reasoning in Tenejapan Mayans. *Cognition, 120*, 33–53. doi:10.1016/j.cognition.2011.02.012
- Li, P., & Gleitman, L. R. (2002). Turning the tables: Language and spatial reasoning. *Cognition, 83*, 265–294. doi:10.1016/S0010-0277(02)00009-4
- Li, P., Ogura, T., Barner, D., Yang, S. J., & Carey, S. (2009). Does the conceptual distinction between singular and plural sets depend on language? *Developmental Psychology, 45*, 1644–1653. doi:10.1037/a0015553
- Macmillan, N., & Creelman, C. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- MacNeilage, P. F., & Davis, B. L. (2000, April 21). On the origin of internal structure of word forms. *Science, 288*, 527–531. doi:10.1126/science.288.5465.527
- McAllister, T. W., Saykin, A. J., Flashman, L. A., Sparling, M. B., Johnson, S. C., Guerin, S. J., . . . Yanofsky, N. (1999). Brain activation during working memory 1 month after mild traumatic brain injury: A functional MRI study. *Neurology, 53*, 1300–1308. doi:10.1212/WNL.53.6.1300
- New, B., Brysbaert, M., Segui, J., Ferrand, L., & Rastle, K. (2004). The processing of singular and plural nouns in French and English. *Journal of Memory and Language, 51*, 568–585. doi:10.1016/j.jml.2004.06.010
- Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Science, 12*, 237–241. doi:S1364-6613(08)00113-710.1016/j.tics.2008.02.014
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108*, 155–184. doi:10.1016/j.cognition.2008.02.007
- Papafragou, A., Massey, C., & Gleitman, L. R. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition, 84*, 189–219. doi:10.1016/S0010-0277(02)00046-X
- Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Science, 13*, 439–446. doi:10.1016/j.tics.2009.07.001
- Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129*, 369–398. doi:10.1037/0096-3445.129.3.369
- Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21*, 150–164. doi:10.1016/S0022-5371(82)90521-7
- Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the investigation of language and thought* (pp. 157–191). Cambridge, MA: MIT Press.
- Stevens, K. (1989). On the quantal nature of speech. *Journal of Phonetics, 17*, 3–45.
- Whorf, B. L. (1956). *Language, thought and reality*. Cambridge, MA: MIT Press.
- Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*, 253–265. doi:10.1002/wcs.104
- Yuzawa, M. (2001). Effects of word length on young children's memory performance. *Memory & Cognition, 29*, 557–564. doi:10.3758/BF03200457
- Zapf, J. A., & Smith, L. B. (2008). Meaning matters in children's plural productions. *Cognition, 108*, 466–476. doi:10.1016/j.cognition.2008.03.008
- Zapf, J. A., & Smith, L. B. (2009). Knowing more than one can say: The early regular plural. *Journal of Child Language, 36*, 1145–1155. doi:10.1017/S0305000909009374
- Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.

Received October 16, 2012

Revision received February 5, 2013

Accepted April 11, 2013 ■