Monolinguals' and Bilinguals' Use of Language in Forming Novel Object Categories

Sarah Fairchild (SarahF@Udel.Edu)
Anna Papafragou (APapafragou@Psych.Udel.Edu)
Department of Psychological & Brain Sciences, University of Delaware
105 The Green, Newark, DE 19716 USA

Abstract

Monolinguals and bilinguals differ along a number of dimensions, including way they label existing object categories (Pavlenko & Malt, 2011). In the present study, we ask whether English monolinguals, Spanish-English bilinguals, and English-Spanish bilinguals also differ in the way they use language when forming novel categories. Previous research with monolinguals shows that a shared label encourages children (e.g., Waxman & Markow, 1995) and adults (e.g., Lupyan, Rakison, & McClelland, 2007) to place objects together. Our results further demonstrate that when two objects shared a Licit Word label like "zeg," monolinguals and bilinguals alike are encouraged to group them together. Illicit Words like "gxz," on the other hand, only influence the categorization decisions of bilinguals. Thus, bilinguals appear to be more flexible in their use of linguistic information in categorization. Neither group made use of non-linguistic cues (patterned frames), suggesting a unique role for language in category formation.

Keywords: verbal labels; categorization; bilingualism

Introduction

The uniquely human faculty of language enables us to convey incredibly complex information to one another. However, language has more to offer than its primary function as a communication tool. Language can be used as a cognitive tool to aid in resource-intensive operations, such as spatial (Loewenstein & Gentner, 2005) and numerical reasoning (Dehaene et al., 1999). Some argue that language is a general computational tool, allowing us to simplify and more efficiently work with information in our environment through a recoding process (Clark, 1998). Here, we investigate how language specifically aids the formation of novel categories by highlighting deep conceptual relationships among objects.

People are more likely to believe that two objects belong together when they share a label (Gelman & Markman, 1986; Jones, Smith, & Landau, 1991; Waxman & Markow, 1995; Balaban & Waxman, 1997; Waxman, 1999; Nazzi & Gopnik, 2001; Sloutsky, Lo, & Fisher, 2001; Booth & Waxman, 2002b; Fulkerson & Waxman, 2007; Lupyan et al., 2007; Diesendruck & Peretz, 2013; Johanson & Papafragou, in press). For example, adults can learn a labeled category of 'aliens' faster than an unlabeled category (Lupyan et al., 2007). Likewise, labels guide children and adults in forming novel plant and animal categories when visual cues are ambiguous (Johanson & Papafragou, 2016). In this sense, language acts as a tool to form categories of novel entities (Wolff, 2011). Importantly, previous research suggests that non-linguistic cues are not as powerful as linguistic labels in influencing categorization. Boutonnet and Lupyan (2015) found that participants were faster to respond to label-object pairs (the word 'dog' – picture of a dog) than sound-object pairs (a dog bark – picture of a dog; cf. also. Lupyan and Thompson-Schill, 2012). Furthermore, linguistic labels are more effective than numbers or symbols in facilitating novel category formation (Gervits et al., submitted). Even for young infants, tones (Ferry, Hespos, & Waxman, 2010), content-filtered speech (Balaban & Waxman, 1997), and primate vocalizations (Ferry, Hespos, & Waxman, 2013) are less effective than labels in facilitating categorization. While it is clear that language can play an important role in category formation, to our knowledge all research to date has been conducted with monolingual participants. However, there are reasons to believe that bilinguals may differ from their monolingual counterparts in their use of labels in categorization. First, monolinguals and bilinguals naturally differ in terms of language abilities and experience. The regular use of two languages changes the way a person's first language is processed (e.g., Van Hell & Dijkstra, 2002) and may even lead to cognitive differences (e.g., Prior & MacWhinney, 2010). Second, we know that bilinguals describe existing categories differently from monolinguals in both their first and their second languages (Pavlenko & Malt, 2010). Additionally, bilingual children are more flexible than monolingual children in the way they apply labels to novel objects (they are less likely to apply the mutual exclusivity constraint when acquiring new words; Davidson, Jergovic, Imami, & Theodos, 1997), and in adulthood, bilinguals are better at acquiring novel labels (Kaushanskaya & Marian, 2009). It may be the case, then, that due to their specific linguistic experiences and advantages, language is a more available and/or flexible tool for bilinguals. For purposes of categorization, bilinguals may weigh linguistic labels more heavily than monolinguals; alternatively, or additionally, bilinguals may be more flexible in regards to what can be considered a label. In either case, bilinguals would not simply be "two monolinguals in one" (Grosjean, 1989) in terms of their use of language as a tool for categorization.

The Present Study

In the present study, we aim to fill a gap in the existing literature by investigating how monolingual and bilingual adults use labels to form new categories. Our goal is to determine to what extent well-formed labels are unique and

powerful category markers for monolinguals and bilinguals alike, as compared to other salient linguistic and non-linguistic cues. Unlike past studies, we take steps to ensure that all types of cues are equally discriminable and thus that any advantage of labels cannot be attributed to superficial features of the cues themselves.

Specifically, our experiment tests the effectiveness of the following cues on the categorization of novel objects: (a) novel Licit Words like "zeg" and "wob," (b) Illicit Words such as "xvxs" that possess linguistic features but are not phonologically possible labels (and as such resemble reverse speech or 'content-filtered speech' – see Balaban and Waxman, 1997), and (c) patterned Frames that may be used to indicate some degree of similarity between the objects contained in them. Following prior studies (Gelman & Markman, 1986; Sloutsky et al., 2001; Diesendruck & Peretz, 2013; Johanson & Papafragou, 2016, among others), we adopt a strong test of the effect of such cues on categorization by asking whether a shared Licit Word, Illicit Word or Frame can lead participants to group together objects that vary in perceptual similarity to each other, from extremely visually dissimilar to nearly identical.

Of interest is whether Licit Words are a more powerful cue compared to Illicit Words and Frames for both monolingual and bilingual speakers – especially for perceptually dissimilar Standard/Target pairings – and if so, whether the role of these cues is comparable across monolinguals and bilinguals. One prediction that one might make is that because language is a more available tool for bilinguals, they will be more influenced by the Licit Words. Alternatively, due to experience with two different linguistic systems, bilinguals may be more likely than monolinguals to make use of the Illicit Words. We do not predict any group differences in the use of the non-linguistic Patterned Frames. To further specify any potential effects of bilingualism, we compare bilinguals with a lifetime of experience using two languages and bilinguals who acquired a second language later in life.

Method

Participants

Thirty-two English monolingual adults (19 female) aged 18-21 (M = 18.53, SD = .76) and thirty-two English-dominant bilingual adults (23 female) aged 18-21 (M = 19.03, SD = .97) participated. In the bilingual group, there were sixteen Spanish-English bilinguals who learned Spanish as a native language but learned English in early childhood (before age 5) and spoke English in educational and professional contexts (i.e., heritage language speakers). The other sixteen participants in that group were English-Spanish bilinguals, who learned English as a native language and Spanish as a second language (L2) later in life (i.e., adult L2 learners). All reported English as being their dominant language, and the groups did not differ in terms of self-rated English proficiency. As might be expected, Spanish-English bilinguals rated themselves as having significantly higher Spanish proficiency than the English-Spanish bilinguals, $t(30) = 3.175$, $p = .003$. Bilingual participant characteristics are presented in Table 1. All monolingual and bilingual participants were undergraduate students at the University of Delaware, and received course credit for their participation.

Bilingual Type	Gender	Age	English Prof.	Spanish Prof.
English-Spanish	3 Male 13 Female	18.94 *1*	9.84 *.35*	6.84 *2.08*
Spanish-English	6 Male 10 Female	19.13 *.96*	9.42 *1.1*	8.64 *.88*
Total	9 Male 23 Female	19.03 *.97*	9.63 *.83*	7.74 *1.82*

Table 1: Bilingual participant characteristics (*SD* in italics).

Materials

Thirty-two grayscale photographs of objects chosen to be unfamiliar to participants were used as stimuli. These objects included arcane tools and pieces of other man-made devices. To ensure novelty, participants were asked at the completion of the experiment if they recognized any of the objects. Only one participant correctly identified a strawberry huller.

The 32 novel objects were then divided into 16 pairs, with one object in the pair designated as Standard A, and the other as Standard B. Each pair was morphed together using the Fantamorph program. For each pair, five morphed pictures (Targets) were created at 10%, 30%, 50%, 70%, and 90% similarity to Standard A (according to program specifications). See Fig. 1 for an example.

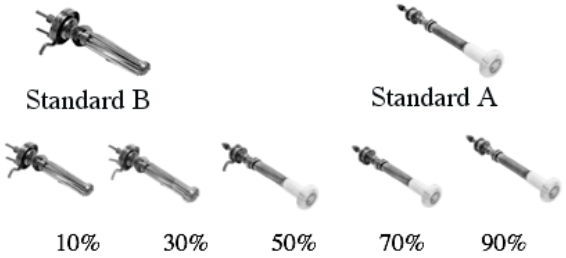

Figure 1: Example stimuli. Percentages reflect target similarity to Standard A.

Eight novel Licit Words were generated using the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002). The Licit Words were between three and five letters in length, and consisted of only orthographically existing onsets, orthographically existing bodies, and legal bigrams (e.g., "zeg"). Additionally, all Licit Words were structured such that they could reasonably be English nouns. Eight Illicit Words that could not exist as English (or Spanish) words were also created. These Illicit Words were designed to be unpronounceable to English speakers, and were created by taking the reverse of the novel Licit Word and replacing the vowels with x's. For example, "zeg" would

become "gxz." Finally, eight different patterned Frames (e.g., solid, dashed) consisting of various geometric shapes (e.g., rectangle, hexagon) were created. Each of these frames was large enough to surround a Target or Standard.

Norming. Two norming studies were conducted to ensure that the stimuli were well-controlled. First, it was important to assess whether or not our participants would be sensitive to the similarity between Standards and Targets produced by morphing. A separate group of 10 monolingual English-speaking adults were presented with all possible combinations of each pair of Standards with a corresponding Target, and were asked to rate the similarity of the Target to one of the Standards on a 9-point scale, 10-90%. Participants were very accurate at rating the 30%, 50%, and 70% morphs; ratings did not differ significantly from actual similarity (all p's < .05). Ratings for the 10% and 90% Target stimuli differed significantly from actual similarity, but in the predicted direction; additionally, 10% objects were rated significantly lower than 30% objects, $t(9) = 7.003$, $p < .001$, and 90% objects were rated significantly higher than 70% objects, $t(9) = 10.902$, $p < .001$.

Second, we ensured that all three cue types were equally discriminable. A separate group of 10 monolingual adults from the same population were asked to rate the discriminability of pairs of Licit Words, Illicit Words, and Frames. These pairs are associated with pairs of Standards in the experiment. For each pair, they answered the question 'How similar are these two items?' on a seven-point scale. Ratings did not differ across cue types (p's > .1).

Procedure

Participants were tested individually or in pairs in a quiet room. All tasks were administered on a laptop with a 15.4" screen, approximately 24 inches away from the participant. The session lasted approximately one hour. Before the experiment began, participants completed a language history questionnaire to assess self-rated English (and Spanish, for bilingual participants) speaking, reading, and writing proficiency (from 1-10, 10 being the highest). All participants were tested in English.

Participants were presented with displays that contained triads of objects: Standards A and B from each of the 16 pairs of novel objects and one of the corresponding Targets. Within each triad, the Standards were presented at the top of the screen and the target was centered at the bottom of the screen, separated by a visible line from the Standards (see Fig.2). All three objects in the triad appeared at once and remained on the screen for 1 second. At that point, either one type of cue was introduced (Licit Words, Illicit Words, or Frames condition), or the display remained the same (No Cue condition). Licit and Illicit Words appeared below the objects and Frames surrounded the objects in a triad (see Fig.2). If a cue was present, the cue for the Target always matched the cue for Standard A. After 2 seconds, any cues disappeared and a red frame appeared around the whole display.

At the beginning of the task participants were given on-screen instructions that were reiterated verbally by the experimenter. They were told that they would see groups of three objects, and that they would be given time to look at the objects. Then, a red frame would appear on the screen. When the red frame appeared, they should press a key to indicate where they think the bottom object belongs, either with the object on the right or the object on the left. Participants were instructed to make a response as quickly as possible.

There were four stimuli blocks within the experiment, one for each cue type, counterbalanced across participants, and stimuli were fully rotated. Within each block, participants saw all Target/Standard triads in a random order, 20 trials per block, for a total of 80 trials. Assignment of Standard A to an object within a triad, as well the left/right position of Standard A within a display, was also counterbalanced. The task was administered using the OpenSesame experimental presentation software (Mathôt, Schreij, & Theeuwes, 2012).

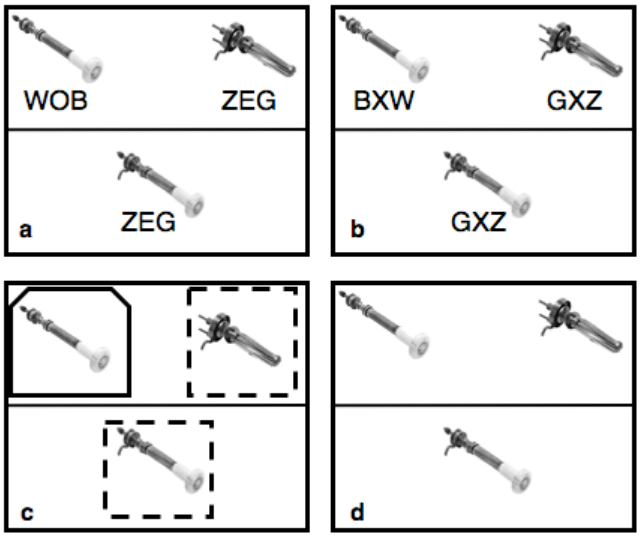

Figure 2: Example stimuli for (a) Licit Word, (b) Illicit Word. (c) Frame, and (d) No Cue blocks.

Following the categorization task, participants completed a lexical decision task as a measure of English proficiency. This task is widely used in both first and second language research, and has been shown to accurately predict language proficiency in various groups of monolinguals and bilinguals (Harrington, 2006), making it an excellent proxy for English language proficiency for our three participant groups. Twenty English words and twenty pseudowords were presented in a random order for 500 milliseconds each. Participants were asked to indicate by keypress, as quickly as possible, whether or not each stimulus was a real English word. Both accuracy and reaction times were collected. For the analysis of reaction times, incorrect responses and trials over 3000 milliseconds were excluded. This task was used to control for English proficiency when performing direct comparisons between the groups.

Results and Discussion

Monolingual Group

A 4 (Cue Type: Licit Word, Illicit Word, Frame, No Cue) by 5 (Perceptual Similarity: 10%, 30%, 50%, 70%, 90%) repeated measures ANOVA was performed with cue compliance as the dependent variable. We considered cue compliant responses to be those where the participant categorized the Target object with the cue-matched Standard A. Cue compliance in the No Cue block was calculated in the same way (i.e., in terms of responses to the predetermined Standard A).

The ANOVA yielded significant main effects of Cue Type, $F(3, 93) = 9.498$, $p < .001$, $\eta^2 = .063$, and Perceptual Similarity, $F(2.44, 75.64) = 709.086$, $p < .001$, $\eta^2 = .823$, as well as a significant interaction between the two, $F(6.14, 190.46) = 2.911$, $p = .001$, $\eta^2 = .043$. As expected, cue compliance increased as Perceptual Similarity increased. Furthermore, there was significantly greater cue compliance in the Licit Word condition ($M = .66$, $SD = .12$) as compared to the No Cue control ($M = .54$, $SD = .06$), $F(1, 31) = 24.144$, $p < .001$, $\eta^2 = .438$, but no significant difference between the Illicit Word ($M = .58$, $SD = .10$) or Frame ($M = .57$, $SD = .11$) conditions and the No Cue condition (both p's $> .1$). More specifically, in the 10% similarity condition, cue compliance in the Licit Word condition was significantly higher than in the Illicit Word ($p = .046$), Frame ($p = .005$), and No Cue ($p = .003$) conditions. The pattern continues into the 30% similarity condition, with higher cue compliance observed for Licit Words as compared to Illicit words ($p = .008$), Frames ($p < .001$), and the No Cue control ($p < .001$). Cue compliance was also significantly higher for the Licit Word condition as compared to the No Cue condition on 50% similarity trials ($p < .001$). No such differences were observed on the 70% and 90% trials, for which cue compliance in the No Cue condition was already at ceiling ($M = .95$ and $M = 1$, respectively). Results are shown in Figure 3.

These findings lead to three main conclusions. First, Licit Words – but not Illicit Words – were treated as labels by our participants. Recall that the cues were presented in a neutral fashion, without any explicit instruction or acknowledgment of their role or significance. Despite this fact, participants spontaneously recruited Licit Words like "zeg" to some degree during their categorization decisions but were not influenced by phonologically impossible Illicit Words like "gxz." Thus, only possible labels in the participant's native language is sufficient to facilitate categorization for monolingual adults.

Second, and perhaps most importantly, we can conclude that language is more powerful than both salient, equally discriminable linguistic (Illicit Words) and non-linguistic (Frames) cues. Labels influence categorization not because they are a simply a salient cue that serves to increase overall similarity, but because adults treat them as category markers.

Finally, we see that participants considered both linguistic and perceptual cues when forming novel categories in our task. When perceptual similarity was high (70% and 90% morphs), participants almost always chose the Standard that shared the same cue as the Target. When perceptual similarity and cues were at odds (10% and 30% morphs), however, participants generally grouped objects that were most similar visually. When the Target was perceptually ambiguous (50%) and could be grouped with either standard, participants were biased towards choosing the label-matched Standard.

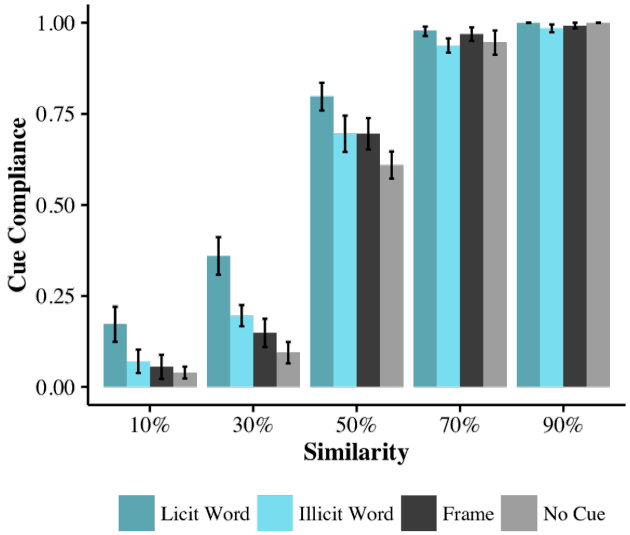

Figure 3: Cue compliance for the monolingual group. Error bars represent +/-1 S.E.M.

Bilingual Group

A mixed ANOVA was conducted with Cue Type (Licit Word, Illicit Word, Frame, No Cue) and Perceptual Similarity (10%, 30%, 50%, 70%, 90%) as within-subjects factors and Bilingual Type (Spanish-English, English-Spanish) as a between-subjects factor. The analysis revealed significant main effects of Cue Type, $F(3, 90) = 11.783$, $p < .001$, $\eta^2 = .081$, and Perceptual Similarity, $F(2.4, 72) = 409.207$, $p < .001$, $\eta^2 = .743$, as well as a significant interaction between the two, $F(6.6, 198) = 5.346$, $p < .001$, $\eta^2 = .069$. As has been observed in the monolingual group, cue compliance increased with Perceptual Similarity, and Licit Word ($M = .70$, $SD = .16$) cue compliance was significantly higher than No Cue ($M = .55$, $SD = .06$) cue compliance, $F(1, 31) = 25.97$, $p < .001$, $\eta^2 = .456$. Unlike the monolingual group, however, there was also significantly higher cue compliance in the Illicit Word condition ($M = .65$, $SD = .15$), as compared to the No Cue condition, $F(1, 31) = 12.130$, $p = .002$, $\eta^2 = .281$. Cue compliance did not differ significantly between the Frame and No Cue conditions, nor between the Licit Word and Illicit Word conditions (p's $> .1$).

Further analyses reveal that for 10% similarity trials, there was higher cue compliance in the Licit Word ($p < .001$) and Illicit Word ($p = .002$) conditions as compared to the No

Cue condition. Cue compliance was also significantly higher for Licit Words ($p < .001$) and Illicit Words ($p = .020$) than Frames in the 10% similarity condition. For 30% similarity trials, cue compliance was significantly higher in the Licit Word ($p < .001$) and Illicit Word ($p = .008$) conditions as compared to the No Cue condition. Furthermore, we observed higher cue compliance for both Licit Words ($p = .006$) and Illicit Words ($p = .020$) in comparison to Frames for 30% similar trials. On 50% similarity trials, again we saw that there was significantly higher cue compliance in the Licit Word ($p < .001$) and Illicit Word ($p = .003$) conditions as compared to the No Cue control. ($p < .001$) Licit Words ($p = .008$), but not Illicit Words ($p > .1$) had significantly higher cue compliance than Frames on 50% similarity trials. There were no differences between cues for the 70% and 90% trials. The main effect of Bilingual Type was not significant, nor were the interactions between Bilingual Type and Cue Type or Bilingual Type and Similarity (all p's > .1). Results from both groups combined are shown in Figure 4.

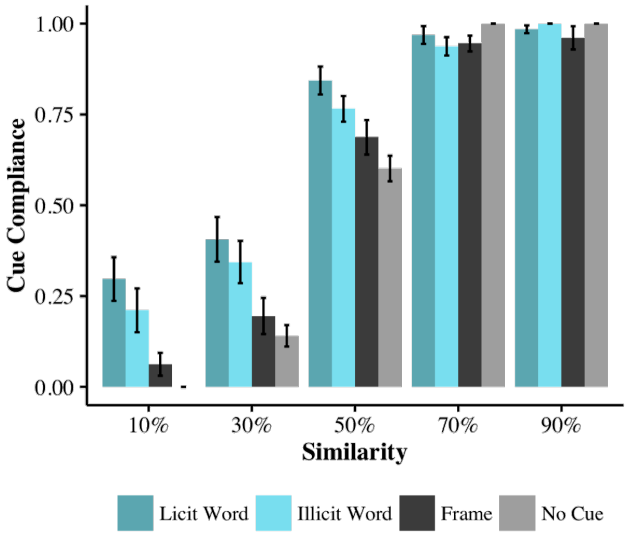

Figure 4: Cue compliance for the bilingual groups (combined). Error bars represent +/-1 S.E.M.

Interestingly, these results indicate that both Licit and Illicit Words are effective in facilitating categorization for bilinguals, and to the same extent. In contrast, only Licit Words influenced the categorization decisions of monolinguals. It is unlikely that the bilingual participants simply were unable to tell the difference between Licit and Illicit words, as they were all English-dominant bilinguals who did not differ from the monolingual participants on a separate test of English proficiency (lexical decision task, not reported here in detail for reasons of space). Rather, what this suggests is that bilinguals are able to use linguistic information more flexibly in category formation, treating linguistic cues that are not possible labels as category markers. Somewhat surprisingly, we also found that there was no difference between Spanish-English bilinguals and English-Spanish bilinguals. Thus, a lifetime of experience using two languages is not necessary to produce this kind of flexibility in the use of linguistic cues for categorization. Instead, a moderately high level of proficiency in a second language appears to be sufficient to cause participants to consider phonologically impossible words as labels. However, this flexibility did not extend to non-linguistic cues. Again, we saw that the patterned frames were not as effective as the linguistic cues, despite the fact that all cues were salient and equally discriminable. In this way, language appears to be a unique tool for promoting categorization. As was observed with the monolingual group, we found that bilinguals were sensitive to both linguistic and perceptual information when making categorization decisions. High visual similarity and a shared linguistic cue (Licit or Illicit Word) encouraged participants to group together two objects.

Conclusion

Previous research indicates that people are more likely to group together objects with a shared label. In the present study, we asked to what extent the influence of labels on categorization is unique for monolingual and bilingual adults by comparing Licit Words to salient, equally discriminable Illicit Words and Frames. Second, we further investigated the influence of bilingual language experience by comparing Spanish-English bilingual heritage speakers to English-Spanish late second language learners.

Our results first suggest that linguistic labels are unique in their ability to facilitate categorization for both monolinguals and bilinguals: two novel objects assigned the same Licit Word were more likely to be grouped together than two objects without linguistic cues. The same effect was not seen for very salient, equally discriminable non-linguistic cues, patterned Frames. Thus, it is not the case that language simply serves to highlight similarities and differences among objects. Rather, linguistic labels (but not shapes) are taken to be category markers. Additionally, we found that perceptual similarity is also taken as an important category marker that works alongside verbal labels. A shared label increased the likelihood of two objects being grouped together, but so did a high level of shared perceptual attributes. These findings contrast with some previous work which has shown that adults completely override perceptual similarity in favor of labels (Sloutsky et al., 2001). However, our results are consistent with more recent work by Gelman and Davidson (2013) which argues for a more nuanced relationship between language and perception in categorization. Because of the influence of similarity, the label advantage is most strikingly observed in the low-similarity conditions (10%, 30%, 50%) for both groups, as cue compliance is already at ceiling for the 70% and 90% similarity conditions.

Critical to our first goal, we found that Spanish-English and English-Spanish bilinguals were influenced not only by Licit Words like "zeg," but also by phonologically impossible Illicit Words like "gxz" when forming novel artifact categories. In this way, bilinguals may be more flexible in the way they integrate linguistic information in

categorization. What we do not know, however, is whether this flexibility comes from differences in the way monolinguals and bilinguals process language, represent object categories (Pavlenko & Malt, 2010), or learn novel words (Kaushanskaya & Marian, 2009).

Interestingly, we found that lifelong experience with two languages is not necessary for this flexibility to arise. We found no differences in behavior between Spanish-English bilinguals who grew up speaking both languages and English-Spanish bilinguals who learned a second language later in life. The flexibility observed here seems to stem, then, from a relatively high level of proficiency in two languages, regardless of age of acquisition.

In sum, we contribute to the existing literature by showing evidence that language is a unique tool for categorization that both monolinguals and bilinguals make use of. Bilingual language experience contributes to increased flexibility in the type of linguistic information that can be considered as a label, but further research is needed to understand the underlying mechanisms and test the limits of this flexibility with groups of bilinguals at varying levels of proficiency.

Acknowledgments

We thank Amelia Weiss for assistance with stimuli creation and data collection.

References

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology, 64*, 3–26.

Booth, A. E., & Waxman, S. (2002b). Object names and object functions serve as cues to categories for infants. *Developmental psychology*, *38*(6), 948.

Boutonnet, B., & Lupyan, G. (2015). Words Jump-Start Vision: A Label Advantage in Object Recognition. *The Journal of Neuroscience, 35,* 9329-9335.

Clark, A. (1998). Magic words: How language augments human computation. *Language and thought: Interdisciplinary themes*, 162-183.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*(5416), 970-974.

Dessalegn, B., & Landau, B. (2013). Interaction between language and vision: It's momentary, abstract, and it develops. *Cognition*, *127*(3), 331-344.

Diesendruck, G., & Peretz, S. (2013). Domain differences in the weights of perceptual and conceptual information in children's categorization. *Developmental Psychology*, *49*, 2383–95.

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3- and 4-month-old infants: an advantage of words over tones. *Child Development*, *81*, 472–9.

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2013). Nonhuman primate vocalizations support categorization in very young human infants. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 15231–5.

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition, 105,* 218-228.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183-209.

Gervits, F., Johanson, M., & Papafragou, A. (submitted). Intentionality and the role of labels in categorization.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and language*, *36*(1), 3-15.

Johanson, M., & Papafragou, A. (2016). The influence of labels and facts on children's and adults' categorization. *Journal of Experimental Child Psychology, 144,* 130-151.

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child development*, *62*(3), 499-516.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*(4), 315-353.

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General, 141,* 170.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking redundant labels facilitate learning of novel categories. *Psychological Science, 18,* 1077-1083.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44,* 314-324.

Nazzi, T., & Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: when does language become a tool for categorization?. *Cognition*, *80*(3), B11-B20.

Pavlenko, A., & Malt, B. C. (2010). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition*, *14*(01), 19-45.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55A*, 1339-1362.

Sloutsky, V. M., Lo, Y. F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development*, *72*, 1695-1709.

Waxman Sandra, R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive Psychology*, *29*, 257-302.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 253–265.