

Simulating Bilingual Word Learning: Monolingual and Bilingual Adults' Use of Cross-Situational Statistics

Kin Chung Jacky Chan (k.c.chan@lancaster.ac.uk)

Lancaster University, Department of Psychology, Lancaster, LA1 4YF, UK

Padraic Monaghan (p.j.monaghan@uva.nl)

University of Amsterdam, Faculty of Humanities, Amsterdam, 1000 BP, NL

Lancaster University, Department of Psychology, Lancaster, LA1 4YF, UK

Abstract

Children learning language in multilingual settings have to learn that objects take different labels within each different language to which they are exposed. Previous research has shown that adults can learn one-to-one and two-to-one word-object mappings via cross-situational statistical learning (CSSL), and that socio-pragmatic cues may differentially influence monolingual and bilingual adults' learning of such mappings. However, the extent to which monolingual and bilingual learners can keep track of multiple labels from multiple speakers has not yet been investigated. We manipulated the number of speakers in a CSSL task that involved learning both mapping types. We successfully replicated previous studies that found that both monolinguals and bilinguals could learn both types of mappings via CSSL. In addition, we found that bilinguals showed a steeper learning rate for two-to-one mappings than monolinguals, and bilinguals were more likely to accept two words for the same object than monolinguals. These results show that the effect of speaker identity on tracking word-object mappings varies according to language experience.

Keywords: statistical learning; bilingualism; mutual exclusivity; cross-situational learning; word learning

The learning of the mapping between a word and its referent is difficult, as there are infinitely many potential referents for a word. This uncertainty is referred to as the "Gavagai" problem (Quine, 1960). The uncertainty is increased still further when children grow up in multilingual environments, as this means there are multiple words for a particular referent. The present study aims to investigate how speaker identity, as a socio-pragmatic cue, impacts on language learning under such conditions of referential and reference uncertainty.

A prominent suggestion as to how language learners overcome the "Gavagai" problem has been that language learners make use of constraints on which mappings can be formed. For instance, the ME constraint suggests that language learners tend to assign only one word to a referent (Markman & Wachtel, 1988). When language learners hear a novel word and see a familiar object, of which they already know the name, and an unfamiliar object, they would, based on ME, pair the novel word with the unfamiliar object. Other constraints include the whole-object assumption and the taxonomic assumption (Markman, 1991; Markman & Hutchinson, 1984; Markman & Wachtel, 1988). Another account of word learning is the socio-pragmatic account,

which suggests that language learners' word learning rely on their socio-cognitive skills and the social cues available in communicative contexts (Tomasello, 2000). This account explains word learning in terms of language learners' ability to actively monitor others' attention (Akhtar & Tomasello, 1996) and intention (Tomasello & Barton, 1994) to discover intended referents of novel words. In general, both of these accounts posit that language learners make use of certain strategies to limit the number of potential referents for a word to help solve the "Gavagai" problem. Yet, constraining the problem space is not the only way to solve the word-learning problem.

Recently, cross-situational statistical learning (CSSL) ability has been proposed as a valuable contributor to word learning. Though the referent of a novel word might be ambiguous within the context of a single learning instance, across multiple learning instances, learners would be able to track the co-occurrences of the novel word and its referent, with which it reliably appears. This statistical information can then help learners to disambiguate which words refer to which referents. Yu and Smith (2007) presented adults with a series of trials containing two to four unfamiliar objects and novel words. Within each trial, the word-object pairings were ambiguous (i.e., novel words were presented in a random order in all trials and there was no correspondence between the order of words and the location of objects on the computer screen), but across trials, with the presentation of different combinations of novel words and their referring objects, the word-object pairings could become apparent. Yu and Smith found that adults could learn the meanings of words via CSSL. This finding has been replicated in various similar studies (e.g., Fitneva & Christiansen, 2011; Hamrick & Rebuschat, 2012; Monaghan & Mattock, 2012; Vouloumanos, 2008).

In these studies, only one-to-one word-object pairs were used. Yet, although learners favour ME (i.e., one-to-one word-referent mappings) when learning the meaning of words, overcoming ME is important for learning categories, homonyms, and synonyms (e.g., Markman & Wachtel, 1988). It is also particularly important for bilinguals as they have to learn translation equivalents (forming many-to-one word-referent mappings; e.g., both "apple" and "manzana" refer to a particular fruit) and interlingual homographs (forming one-to-many word-referent mappings; e.g., "tuna" refers to a kind of fish in English but prickly pear in Spanish).

Ichinco, Frank and Saxe (2009) familiarised and then tested adults on a set of one-to-one word-object pairs. Then, the participants were familiarised to a second set of one-to-one word-object pairs. Some of the pairs in the second set required the remapping of objects or words. Thus, although each set consisted of one-to-one word-object pairs, across the two sets, there was a combination of one-to-one, two-to-one, and one-to-two word-object pairs. The two-to-one and one-to-two word-object pairs were critical for testing whether adults could relax ME during a CSSL task. It was found that the participants were successful in learning the one-to-one word-object pairs and the first mapping of the two-to-one and one-to-two word-object pairs. By contrast, they failed to learn the second mapping of the two-to-one and one-to-two word-object pairs. Ichinco et al. took the results of their study as evidence against a simple associative learning account of word learning.

Yet, Kachergis, Yu and Shiffrin (2009) argued that the results of Ichinco et al.'s (2009) study could be due to a blocking effect, giving rise to the participants favouring the first mapping learnt. Using a similar paradigm to that in Ichinco et al.'s study, Kachergis et al. manipulated the number of occurrences of the second mapping of the word-object pairs. It was found that the extent to which the participants relaxed ME – successful at learning the second mapping of the word-object pairs – was associated with the number of times they had been exposed to the pairs, such that the participants were more likely to relax ME when there was more evidence (i.e., exposure) in the input for the second mapping.

These CSSL studies examined CSSL in a monolingual population. Only a few studies have looked at CSSL in a bilingual population. A study similar to that of Yu and Smith's (2007) by Escudero, Mulak, Fu and Singh (2016) showed that bilingual adults could learn one-to-one word-object pairs via CSSL, outperforming their monolingual counterparts. Another study by Poepsel and Weiss (2016) investigated whether bilingual adults would learn one-to-two word-object pairs better than monolingual adults do, owing to them encountering more instances where they have to relax ME in order to learn new words. They tested the participants' learning of the first and second word-object mappings of the one-to-two word-object pairs in separate testing blocks after the first and second block of learning trials respectively, and tested all word-object mappings in the final testing block after the third learning block. Consistent with Poepsel and Weiss' prediction, it was found that the bilingual adults were quicker than the monolingual adults at learning and showed higher proficiency in learning the one-to-two word-object pairs.

Further, Benitez, Yurovsky and Smith (2016) familiarised monolingual and bilingual adults with a set of one-to-one and two-to-one word-object pairs and tested their learning of the word-object mappings. They found that the monolingual and bilingual adults performed similarly on the task. Both groups showed learning of both the one-to-one and two-to-one word-object pairs, but both groups were better at learning the one-to-one pairs. This is surprising, but not unreasonable, as

monolinguals, who have to learn synonyms, are also experienced in learning two-to-one word-object mappings. An interesting finding of their study was that when a phonological cue distinguished sets of labels, the bilingual adults were more likely to learn both words of the two-to-one pairs. This suggests that bilingual adults are more sensitive to the linguistic cues that hint at different languages present in the linguistic input. Taken together, there is evidence that bilingual adults are better than their monolingual counterparts when it comes to learning word-object pairs that violate ME via CSSL.

Other studies have investigated whether socio-pragmatic cues in the linguistic input would affect learners' cross-situational word learning (e.g., Metzling & Brennan, 2003; Trude & Brown-Schmidt, 2012). Poepsel and Weiss (2014) manipulated the socio-pragmatic information available to participants. In one condition, the participants were told that there were two languages involved in the task. In the other two conditions, the participants were not told anything explicitly, but in one of these conditions, the participants were provided with information on speaker identity – they heard a male and a female voice. In the two-voice condition, the two speakers used the same word to refer to a different object, which could be seen as an implicit cue that there could be two different linguistic structures involved in the task. It was found that the manipulation of socio-pragmatic information did not affect the monolingual adults' performance on learning one-to-two word-object pairs. Yet, in multilingual environments it is more usual for one object to be labelled differently by distinct speakers. Whether varying speaker identity would affect bilingual adults' cross-situational word learning, and whether speaker identity can influence learning of two-to-one mappings is as yet unknown.

The aim of the present study was to examine whether speaker identity would differentially affect monolingual and bilingual adults' performance on a CSSL task that involved the learning of one-to-one and two-to-one word-object pairs. We included two conditions – one where there was a single speaker labelling objects in two ways, and one where different speakers labelled objects in two ways. The present study employed a CSSL paradigm similar to that in Monaghan and Mattock's (2012) study, which is slightly different from many of the CSSL paradigms used in other studies. The crucial difference was that the CSSL paradigm used in the present study did not distinguish between familiarisation and test trials – participants were required to make a forced choice response, without feedback, between two objects in all trials. This allowed an online measure of how quickly and reliably participants form one-to-one and two-to-one word-object mappings across trials.

Another unique feature of the present study was that an additional ME block was administered at the end of the CSSL paradigm to determine whether successful learning of two-to-one word-object pairs was due to successful tracking of two structures in the linguistic input or a general tendency to relax ME.

It was predicted that bilingual adults would be quicker and more accurate at learning two-to-one word-object pairs than monolingual adults. Also, it was predicted that the presence of speaker identity would further benefit bilingual adults' learning of two-to-one word-object pairs due to them being more experienced than monolingual adults in using socio-pragmatic information to track multiple structures in their linguistic input.

Method

Participants

Forty monolingual ($M_{\text{age}} = 22.80$, $SD = 4.56$, 4 male) and forty bilingual ($M_{\text{age}} = 23.58$, $SD = 3.71$, 10 male) participants were recruited through SONA (the departmental online recruitment system) and advertisements on social networking websites. Half of the participants in each language group were randomly assigned to the one-speaker condition, and the other half the two-speaker condition. Nine additional participants were tested but excluded due to technical difficulties ($n = 8$) and experimenter error ($n = 1$).

Participants rated their language proficiency on a 10-point Likert scale from 1 (limited knowledge) to 10 (highly proficient). Monolinguals rated their English proficiency at an average of 9.95 ($SD = 0.22$). Ten monolingual participants indicated exposure to additional languages, but were considered functionally monolingual, as all such proficiency ratings were below 4 ($M = 2.23$, $SD = 0.93$), a similar cut-off to that used in Poepsel and Weiss (2016). The bilingual group rated the proficiency of their first language at an average of 9.85 ($SD = 0.43$) and that of their additional languages at an average of 7.36 ($SD = 2.01$).

Materials and apparatus

Fourteen images of unfamiliar objects and 20 novel words were selected from the Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016). Sound files of the novel words were generated using the system voices Kate (female voice) and Daniel (male voice) on Macintosh computers. Pictures were randomly paired with the novel words for each participant, such that there were eight one-to-one word-object pairs and six two-to-one word-object pairs. In the one-speaker condition, all words were uttered by the same speaker. The gender of the speaker was counterbalanced across participants assigned to the one-speaker condition. In the two-speaker condition, half of the words were uttered by a male, and the other half a female. For words in the two-to-one word-object pairs, the two words referring to the same object were uttered by voices of different gender. The gender of speaker of each word was counterbalanced across participants assigned to the two-speaker condition. In addition, eight images of familiar objects were selected from the TarrLab Object Databank (1996) for use in the familiarisation trials. Sound files of the familiar words were generated using the system voice Allison (female voice) on Macintosh computers. Note that this was a different voice from those used in the test trials, so that the

participants did not have any reliable information on what language(s) the speakers in the test trials spoke. The pictures and audio files of words were presented on a Macintosh computer using PsychoPy (Peirce, 2009).

Procedure

The experiment took place in a quiet room. Participants were tested in groups of less than five people. After receiving an information sheet and signing informed consent, each participant was asked to complete the experiment on a Macintosh computer. Participants were asked to put on headphones for the experiment.

For each trial, the participants saw two pictures presented on the screen. After 500 ms, they heard a word. The target and foil were randomised for screen position (left vs. right) across trials. The participants were instructed to press the right arrow key if they thought the word presented refers to the object on the right and press the left arrow key if they thought the word presented refers to the object on the left. The participants were also instructed to make a guess if they did not know the answer to any of the test trials.

The participants first took part in a familiarisation block, in which they were presented with four trials containing known words and objects. This was to familiarise the participants with the experimental procedure. For the main experiment, the participants first took part in eight CSSL blocks of 40 test trials each. Within each of the CSSL blocks, each object occurred four times as the target and four times as the foil. The screen position of the target and foil were pseudo-randomised, such that the target appeared an equal number of times as the left and as the right object. Words in the one-to-one word-object pairs occurred four times within a block, whereas those in the two-to-one word-object pairs occurred only two times within a block. The order of trials within each block was pseudo-randomised, such that none of the objects appeared in two consecutive trials. An important point to note is that the participants were not provided with any information on the number of languages involved in the main experiment – the only socio-pragmatic cue available to them was the number of speakers in the task. The participants were allowed to take a short break after every two blocks. After all eight blocks, the participants were exposed to each one-to-one word-object pair 32 times and each two-to-one word-object pairs 16 times.

Immediately after the eighth CSSL block, the participants took part in an ME block containing eight test trials. Each trial featured one of the objects from the one-to-one pairs from the CSSL blocks and a new unfamiliar object. Each object occurred one time as the target and one time as the foil. As in the CSSL blocks, the screen positions of the target and foil were pseudo-randomised. For each of the first four trials, the participants heard a word that they had just had the opportunity to learn during the CSSL blocks. These four trials served the purpose of familiarising the participants with the new unfamiliar objects and to control for a possible novelty bias during later trials, where the new unfamiliar objects were the target. For each of the final four trials, the participants

heard a new novel word, which was spoken by the speaker who spoke the word for the foil in the same trial. These final trials were critical for determining the extent to which the participants relied on ME when learning new words. If a participant was relying on ME, they would be more likely to choose the familiar object in the first four trials and the less familiar objects in the last four trials. However, if a participant was relaxing ME, their performance would be at chance level – choosing either object as the answer in any given trial.

Upon completing the ME block, all participants were given a full debrief and received £3.50 for taking part in the experiment. Each testing session lasted less than 30 minutes.

Results

Learning over the training blocks

Data from six participants, one from the monolingual group and five from the bilingual group were excluded from analysis, due to them not demonstrating learning across testing blocks (i.e., average proportion correct across first two blocks > average proportion correct across final two blocks).

To compare whether number of speakers had influenced the monolingual and bilingual adults' learning of the two types of mappings, generalised linear mixed-effects (GLM) modelling was used to predict the adults' response accuracy. The data for GLM modelling consisted of the response accuracy from each participant on each trial, giving a total of 23680 observations.

A series of GLM models were fitted using the `glmer` function (family = binomial) in the `lme4` package in R. A backwards elimination approach was used, entering as fixed factors: language group, speaker number, block, mapping type of the target (whether it had one or two labels), and mapping type of the foil. Extraneous variables, including participant gender and speaker gender, did not influence the participants' performance. For training accuracy, the best model ($AIC = 19977.7$, $BIC = 20131.1$, $logLik = -9969.9$, deviance = 19939.7) given the data is the model with the following fixed effects: the three-way interaction, all two-way interactions, and main effects of block, language group, and target mapping and the main effect of foil mapping; the following random intercepts: subject, word, target, and foil; and the following random slopes: block on subject and language group on word and target.

As expected, there was a significant effect of block ($\beta = 0.26655$, 95% CI [0.2204, 0.3127]), suggesting that, in general, performance improved across testing blocks. The main effect of target mapping was also significant ($\beta = 0.74309$, 95% CI [0.5191, 0.9670]). To our surprise, and contrary to Benitez et al. (2016), the participants were *better* at learning the two-to-one than one-to-one mappings. There was also a significant main effect of foil mapping ($\beta = 0.26248$, 95% CI [0.1826, 0.3424]), suggesting that performance was better if the foil in a given trial was a two-to-one mapping. In addition, the interaction between block and target mapping was also significant ($\beta = 0.06914$, 95%

CI [0.0194, 0.1189]), showing a convergence of the participants' performance in learning the two mapping types across blocks, such that although their learning of the two-to-one mappings was better than that of the one-to-one mappings across blocks, their learning rate for the one-to-one mappings was steeper.

Though there was no significant main effect of language group, the interaction between language group and target mapping was significant ($\beta = -0.49190$, 95% CI [-0.8068, -0.1770]), indicating that although both language groups were better at learning the two-to-one mappings, the monolingual group's performance difference between the two mapping types was greater than that of the bilingual group.

Finally, there was a significant three-way interaction between block, language group, and target mapping ($\beta = 0.08267$, 95% CI [0.0103, 0.1551]; see Figure 1). The three-way interaction suggests that, for the monolingual group, the learning rate of the two-to-one mappings was more gradual than that of the one-to-one mappings, whereas for the bilingual group, there was faster learning of the two-to-one mappings over the blocks.

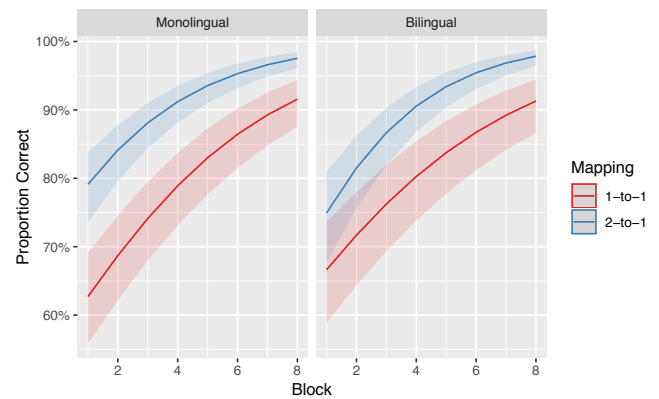


Figure 1: Three-way interaction of block, language group and target mapping.

Performance on the ME task

Though there were no significant main effect or interactions with number of speakers in the task, it was possible that monolingual and bilingual speakers relied on different strategies – either relaxing ME or successfully tracking two labels in the linguistic input would produce a similar pattern of results. In order to determine whether the two language groups relied on similar strategies, their performance in the ME block was analysed. In any given trial, if a participant picked the object that was in line with the application of ME, they scored 1, otherwise they scored 0. Similar to the treatment of the data from the CSSL blocks, GLM models were fitted to participants' scores on each trial (592 observations). Predictor variables of the GLM models were language group, speaker number, and word type (familiar vs. new), and a backwards elimination approach was used.

Extraneous variables, including participant gender and speaker gender, did not influence the participants'

performance. The best model ($AIC = 180.6$, $BIC = 202.6$, $\log Lik = -85.3$, deviance = 170.6) given the data is the model with the following fixed effects: the two-way interaction and the main effects of language group and speaker number; and the random intercept of subject.

Of particular note, the significant interaction between language group and speaker number ($\beta = 3.3748$, 95% CI [0.0054, 6.7442]; see Figure 2) suggests that although both language groups were able to systematically apply ME, there was a tendency for the bilingual group to relax ME when there was only one speaker in the task.

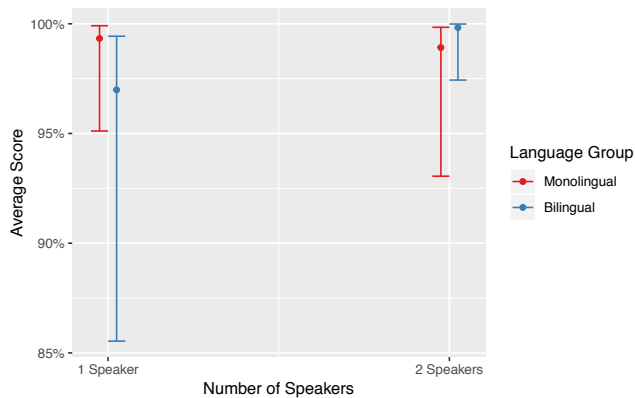


Figure 2: Interaction between speaker number and language group.

Discussion

Using a CSSL paradigm, the influence of speaker identity on monolingual and bilingual adults' learning of one-to-one and two-to-one word-object pairs was examined. In line with previous research (e.g., Benitez et al., 2016), our results showed that both monolingual and bilingual adults are capable of learning one-to-one and two-to-one word-object mappings through CSSL. Yet, inconsistent with Benitez et al.'s main finding, both groups of participants in the present study were better at learning the two-to-one than one-to-one word-object mappings. This could be due to the imbalanced number of objects for each type of mapping – there were six objects that mapped onto two words, whereas only four objects mapped onto one word. As the majority of the objects had two names, it was possible that the participants had formed an expectation that each object could take on two names and used this as a learning strategy for the task.

This explanation is given further weight by the finding that the participants were more likely to accurately map a label to the target object if the foil object was a referent of a two-to-one mapping. The indication here is that, in any given trial, if a participant had already learnt two labels for the foil, they would map the different word presented to the target, due to the foil already taking on the expected maximum number of words. Yet, in order to detect that some objects were named with one, and others with two, labels, participants had to gain this knowledge from tracking implicitly the association between particular words and objects over multiple scenes.

That participants were adept at acquiring both one and two labels for objects so early in training demonstrates the power of this learning mechanism. Yet, it should also be noted in Benitez et al.'s study, there were instances where two-to-one mappings were better learnt than one-to-one mappings in that the presence of a second label seemed to have improved learning of the first label. Our task could be showing a similar advantage.

However, there were subtle differences in the learning trajectory of the monolingual and bilingual speakers in our study. The significant three-way interaction between block, language group, and target mapping shows that the learning of the two-to-one mappings was different for the two language groups. The performance of the monolingual group showed less improvement in learning of the two-to-one mappings, whereas the bilingual group had a steeper learning rate for the two-to-one compared to one-to-one mappings. This steeper learning rate could be due to their experience with language. In bilingual adults' linguistic environment, two-to-one word-object mappings would be more dominant than one-to-one word-object mappings. This experience could have benefited them in learning the two-to-one mappings in the task, which is in line with the finding of Kalashnikova, Mattock and Monaghan's (2015) study that bilingual experience would lead to more flexible use of ME, exhibited by higher tendency to accept lexical overlap.

Alternatively, the observed difference between the two language groups could be due to the monolingual adults displaying an early advantage in learning the two-to-one mappings from the first testing block, whereas the bilingual adults' learning of the two mapping types did not differ until the third testing block. A possible explanation to this initial difference of the learning of the two-to-one mappings could, again, relate to the imbalanced number of objects pertaining to each type of mapping. The imbalanced number may have served as a cue for the monolingual adults to more readily learn two words for one object, which could have been salient because this was inconsistent with their usual experience (i.e., one-to-one word-object mappings being the norm). For the bilingual adults, as they frequently confront two-to-one word-object mappings, the imbalance may be less salient and thus a less effective cue to influence their cross-situational learning early on in the experiment. These significant interactions suggest that language experience plays a role in the application of different word-learning strategies.

However, our results for the bilingual participants do not tally with those of Benitez et al.'s (2016) finding, which showed that bilingual adults' learning of two-to-one word-object mappings in a CSSL task was worse than their learning of one-to-one mappings. In Benitez et al.'s study, participants were presented with four objects and four words at a time during training, whereas the participants in the present study were only presented with two objects and one word at a time. The complexity of Benitez et al.'s task could have favoured the learning of one-to-one mappings. In their study, although the number of co-occurrences of each corresponding word-object pair was the same for both mapping types, the spurious

co-occurrences of unpaired word-object mappings was higher for the two-to-one mappings, making the learning of the two-to-one mappings more difficult than that of the one-to-one mappings. In the present study, although the participants were presented with fewer tokens of the two-to-one mappings, the reduced number of objects and words in a given trial were likely to have more closely mimicked actual word-learning experiences than Benitez et al.'s task, making the learning of both types of mapping relatively easy to the participants in the present study.

In addition, the design of the present study required participants to make a decision about a pairing on every trial, unlike in previous studies where participants went through a familiarization phase and then a test phase. This could have made the participants' learning of the word-object mappings more explicit and highlighted to the participants that the majority of the mappings were two-to-one, giving rise to the observed better learning of the two-to-one mappings. Determining the extent of referential ambiguity and the relative occurrence of two-to-one versus one-to-one mappings in the language learner's experience will enable us to determine more closely which experimental task better resembles natural language learning.

In contrast to our prediction, manipulating speaker identity did not influence CSSL of either language group. It was perhaps less surprising for the monolingual group, as Poepsel and Weiss (2014) found that manipulating speaker identity did not affect monolingual adults' learning of word-object mappings that violate ME. Taking into account Benitez et al.'s (2016) finding that linguistic cue could affect bilingual adults' learning of two-to-one word-object mappings and the non-significant effect of speaker identity in the present study, it is likely that information about the languages involved in the linguistic input per se is more important than speaker identity as a cue in influencing bilingual adults' word learning. In reality, information about languages in the input is a more reliable cue than speaker identity, as one speaker could speak multiple languages and different speakers could speak the same language.

Nevertheless, speaker identity did seem to have an effect on the strategy used by the two language groups. In the ME block, both language groups demonstrated majority use of ME. Yet, when there was only one speaker involved in naming objects, the bilingual group showed a greater tendency to relax ME (Kalashnikova et al., 2015). This suggests that although speaker identity did not have an effect on the observed responses of the participants in the CSSL task, it may have altered the strategies that they use. The bilingual speakers were more likely than the monolingual speakers to relax ME when more than one language structure was used by the same speaker. This may have been due to greater familiarity by bilingual speakers that individuals may speak more than one language.

In a broader sense, the results of the present study have demonstrated that language learners can flexibly use multiple word-learning strategies to learn different language structures in solving the "Gavagai" problem. In an environment with

multiple language structures, learners have to quickly discriminate the different structures (Gebhart, Aslin & Newport, 2009). Previous studies (e.g., Qian, Jaeger & Aslin, 2012) have shown that socio-pragmatic cues, such as a voice change, can help learners focus on the syntactic structures available in the input. The lack of overall influence of speaker identity on the CSSL task in the present study should, therefore, not be taken as evidence that socio-pragmatic cues do not contribute to word learning, as it could instead be that word learning across multiple situations does not rely so heavily on this particular socio-pragmatic cue. Other socio-pragmatic cues, for example information on the languages that the speakers in the CSSL task speak or more information on the speakers' linguistic identities, might be more effective in influencing learners' reliance of word-learning strategies. Nevertheless, the results of the present study, in terms of trajectory of learning on the CSSL task and performance in the ME task, suggest that the extent to which a word learning strategy is relied upon depends in part on an individual learner's previous experience with languages and the learning context. These results also begin to give us some insights into how language experience, contextual cues and task design contribute to shaping learners' use of different word-learning strategies.

In summary, we replicated previous studies that found that language learners are adept at accepting multiple labels for the same object. Curiously, when only one word is heard, and two possible objects viewed, both monolingual and bilingual speakers were better at learning two labels for an object than one label for an object. The effects of participants' linguistic background exerted subtle effects on this ability, with a steeper learning rate of two-to-one mappings for bilinguals compared to monolinguals, and a greater ability for bilinguals to be flexible in the application of the ME constraint. These results show that the parameters determining how word-object mappings are acquired and the role of language experience in driving this learning are complex and varied.

Acknowledgements

This work was supported by a Leverhulme Trust Doctoral Scholarship to KCJC and the International Centre for Language and Communicative Development (LuCiD) at Lancaster University, funded by the Economic and Social Research Council (United Kingdom; ES/L008955/1).

References

- Akhtar, N., & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *Developmental Psychology, 14*, 79-93.
- Benitez, V. L., Yurovsky, D., & Smith, L. B. (2016). Competition between multiple words for a referent in cross-situational word learning. *Journal of Memory and Language, 90*, 31-48.
- Escudero, P., Mulak, K. E., Fu, C. S. L., & Singh, L. (2016). More limitations to monolingualism: Bilinguals outperform monolinguals in implicit word learning. *Frontiers in Psychology, 7*, 1218.

- Fitneva, S. A., & Christiansen, M. H. (2011). Looking in the wrong direction correlates with more accurate word learning. *Cognitive Science*, *35*, 367-380.
- Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science*, *33*, 1087-1116.
- Hamrick, P., & Rebuschat, P. (2012). How implicit is statistical learning?. In P. Rebuschat & J. N. Williams (Eds.), *Statistical learning and language acquisition* (pp. 365-382). Berlin: de Gruyter Mouton.
- Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Unpublished manuscript*.
- Ichinco, D., Frank, M., & Saxe, R. (2009). Cross-situational word learning respects mutual exclusivity. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (pp. 2214-2219). Austin, TX: Cognitive Science Society.
- Kachergis, G., Yu, C., & Shiffrin, R. M. (2009). Frequency and contextual diversity effects in cross-situational word learning. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (pp. 2220-2225). Austin, TX: Cognitive Science Society.
- Kalashnikova, M., Mattock, K., & Monaghan, P. (2015). The effects of linguistic experience on the flexible use of mutual exclusivity in word learning. *Bilingualism: Language and Cognition*, *18*, 626-638.
- Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. P. Brynes, *Perspectives on language and thought: Interrelations in development* (pp.72-106). New York, NY: Cambridge University Press.
- Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*, 1-27.
- Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*, 121-157.
- Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, *49*, 201-213.
- Monaghan, P., & Mattock, K. (2012). Integrating constraints for learning word-referent mappings. *Cognition*, *123*, 133-143.
- Pierce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroscience Methods*, *162*, 8-13.
- Poepsel, T. J., & Weiss, D. J. (2014). Context influences conscious appraisal of cross situational statistical learning. *Frontiers in Psychology*, *5*, 691.
- Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, *152*, 9-19.
- Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to represent a multi-context environment: More than detecting changes. *Frontiers in Psychology*, *3*, 228.
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
- The Object Databank [Computer software]. (1996). Carnegie Mellon University. Retrieved from: <http://wiki.cnbc.cmu.edu/Objects>
- Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics*, *10*, 401-413.
- Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, *30*, 639.
- Trude, A. M., & Brown-Schmidt, S. (2012). Talker-specific perceptual adaptation during online speech perception. *Language and Cognitive Processes*, *27*, 979-1001.
- Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, *107*, 729-742.
- Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*, 414-420.