Title
Social influences on the regularization of unpredictable linguistic variation

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 36(36)

ISSN
1069-7977

Authors
Feher, Olga
Kirby, Simon
Smith, Kenny

Publication Date
2014

Peer reviewed

Social influences on the regularization of unpredictable linguistic variation

Olga Feher (olga.feher@ed.ac.uk), Simon Kirby & Kenny Smith
Language Evolution and Computation Research Unit, School of Philosophy, Psychology & Language Sciences,
University of Edinburgh, Dugald Stewart Building, 3 Charles Street, Edinburgh, EH8 9AD, UK

Abstract

Languages tend not to exhibit unpredictable variation, and learners receiving variable linguistic input tend to eliminate it, making the language more regular. We explore how this behavior is influenced by social cues, in particular when variability is distributed within and across teachers. We trained participants on an artificial language that contained lexical variability and manipulated how that variability was distributed across teachers: learners either received input from one or three variable teachers, or from three teachers who were individually consistent but exhibited variability collectively. We found that learners were more likely to produce variable output when their input came from (one or multiple) teachers who exhibited variable labeling, and they regularized more when learning from individually consistent teachers. This indicates that the propensity of language learners to eliminate linguistic variation is modulated by social cues, pointing to potential links with the broader literature on social learning in other domains.

Keywords: language, unpredictable variation, probability matching, regularization, conformity.

Introduction

A salient feature of natural languages is that they exhibit very little free variation: no two linguistic forms will occur in precisely the same environments and perform precisely the same functions; rather, usage is conditioned in accordance with phonological, semantic, pragmatic or sociolinguistic criteria (Givón, 1985). Where variation does exist, its use is conditioned on some contextual or grammatical variable such as the social situation (Labov, 1963). Pidgin languages form an exception, because adult second language learners tend to be variable (Johnson et al 1996). During creolization (the process where pidgins are learned by children as a first language), however, young learners do not reproduce the variable input in their parents' language, converging instead on more systematic usage (Sankoff, 1979). The same happens when deaf children learn sign language from parents who are non-native, and therefore variable, signers (Singleton & Newport, 2004). This suggests that child learners are biased against unpredictable linguistic variation, and that these biases in child learners reshape languages. Unpredictable variation therefore provides a useful test case for studying biases in language learning and how those biases drive language change, creolization and, ultimately, language design.

Learner responses to unpredictable variation have also been studied under more controlled circumstances, using artificial language learning paradigms. Hudson Kam & Newport (2005) asked adults and children to learn an artificial language containing inconsistent grammatical morphemes, and showed that, while adults tend to probability match the variability present in their input (i.e. producing a variable particle in the proportion it occurred in their input data), children regularize, producing only one (typically the more common) variable. This is consistent with the idea that young learners drive creolization and grammaticalization in the formation of languages. However, other studies have shown that adults will also regularize inconsistent linguistic input under the right conditions, such as when presented with a high level of inconsistency and one determiner that is used much more than the others (Hudson Kam and Newport, 2009), when tested on novel words (Wonnacott and Newport, 2005), or when the language is passed between learners in an iterated learning paradigm (Reali and Griffiths, 2009; Smith & Wonnacott, 2010).

Existing artificial language learning work in this area presents learners with linguistic data produced by a single variable teacher – either the experimenter in face-to-face teaching, or a single voice in an automated training procedure. Learning a natural language typically involves exposure to the linguistic behavior of multiple individuals (parents, siblings, peers, etc). In this paper we therefore explore how learners respond to variability distributed across multiple teachers. When learners are presented with inconsistent information from different teachers, they could acquire all the variants and use them probabilistically (i.e. probability match, as adults tend to in the studies reviewed above) or preferentially imitate the majority behavior, known as conformist copying. Conformist copying has been shown to be an important social learning mechanism in humans (Morgan et al, 2012) as well as in several animal species (Van de Waal et al, 2013, Pike & Laland, 2010). However, whether such social learning strategies play a role in the learning of variable linguistic input remains an open question. We used a simple artificial language learning paradigm (similar to that employed in Reali and Griffiths, 2009, Exp 1, but with participants providing spoken responses) to investigate learners' responses when presented with unpredictable variation of lexical items within and between teachers. We conducted two experiments that investigated different aspects of learning from multiple teachers. Experiment 1 contrasts learning from one variable teacher to learning from several variable teachers, to test whether learning from multiple teachers has any impact on probability matching behavior normally observed in such paradigms. Experiment 2 contrasts variation distributed within teachers with variation distributed across teachers.

Methods

The general procedure was identical across experiments. Participants progressed through a self-paced computer program individually, in sound-proofed booths. Participants underwent a two-stage training and testing procedure on a language which provided 6 artificial words labeling 3 objects, each object with 2 alternative names. We used pre-recorded auditory stimuli for training, and participants provided oral responses during testing.

Participants

Forty-eight native English speakers were recruited through the Student and Graduate Employment job search web site at the University of Edinburgh. Participants were randomly assigned to one of three experimental conditions, with 16 participants per condition, and were paid £8/hour pro rata.

Stimuli

Objects The visual stimuli were colored vintage drawings of three relatively obscure plants: an agueroot, a bilberry and a nightshade (Fig.1).

Figure 1: Drawings of the three plants (from left to right: agueroot, bilberry, nightshade) used as the visual stimuli.

Labels The miniature language consisted of six 3-syllable artificial object labels: *drambowit* (dræmboːwɪt), *elbanage* (ɛlːbənɪdʒ), *stanishote* (stæːnɪʃot), *timplika* (tɪmpliːka), *vamola* (vəmoːlə), *zellerand* (zɛːlərænd). The words were constructed using the ARC Nonword database (Rastle et al. 2002) and conformed to English phonology. A full description consisted of a carrier phrase followed by an object label. We used two possible carrier phrases: *osen pilt* (oːsən pɪlt) and *sholar zene* (ʃoːlər ziːn) chosen randomly for each participant but balanced across participants and conditions.

Speakers We recorded three speakers (A-C) speaking all 12 possible combinations of carrier phrase and object label. All speakers were female, of English origin, and had easily distinguishable accents: one had a Received Pronunciation accent (characteristic of middle class individuals from the South of England), one had a Geordie accent (originating in Newcastle-upon-Tyne in Northern England) and one had a distinctive combination of RP and Standard Scottish English. We obtained 10 recordings of each description from each speaker: during training, on each trial one of these 10 sound files was selected at random, allowing natural phonetic variation in our participants' training data.

Procedure

Training The assignment of labels to objects was independently randomized for every participant, with two randomly-selected labels being associated with each object. For each participant, we also selected a single carrier phrase at random, which was used throughout their training.

There were 9 training blocks, each consisting of 27 trials. In a block, every object was presented 9 times: 6 times with one label, 3 times with another. The order of presentation of the 27 trials was fully randomized within each block. In each trial, the picture of the object appeared on the screen for 1s, then a description was played for the participant through headphones. Participants were then prompted to repeat the sentence orally, and their responses were recorded via a microphone situated in the booth. Repetition of the stimulus sentences was intended to encourage participants to practice the pronunciation of the unfamiliar vocabulary, as well as to ensure they were paying attention to the training input. After repeating the description, participants progressed to the next trial by pressing the space key, or after 10s elapsed, whichever came first. There were no breaks between the training blocks, but after the first four blocks and again before the testing phase, participants were given the option to take a break.

Testing There were 36 testing blocks, each consisting of 3 trials, one for each object. During each trial participants were presented with an image of one of the 3 plants on the screen and prompted to name plant. Their answers were recorded via the microphone, as before. Participants progressed to the next trial by pressing space, or when 9s had elapsed, whichever came first.

Experiment 1

In Experiment 1, there were two experimental conditions (N=16 in each), each providing different social information to the participants. In the *One Speaker* condition, for each participant, one of the three speakers (A, B or C) was randomly selected, and the participant only heard recordings from this speaker. In each training block, the speaker named each object 9 times using one of the labels 6 times (Majority Label) and the other 3 times (Minority Label). This meant that there was unpredictable variation in the single teacher's lexical input, in the ratio 2:1.

In the *Within Speakers* condition, participants heard all three speakers (A, B and C) name objects in the training phase: each speaker named each object 3 times within a single block of training, using both the Majority and Minority Labels in the ratio 2:1. Note that the only difference between conditions was that the participants heard 1 or 3 speakers: the total number of trials and the number and proportion of the Majority and Minority labels was kept constant across conditions.

Participants' responses were transcribed by hand and classified, for each trial, as featuring either the Majority Label, Minority Label or NA (if participants gave no response, used an unintelligible label, or used an illegal label, e.g. a label that in the training phase was used to describe another object).

Results of Experiment 1

Participants' productions during testing in the One Speaker and Within Speakers conditions were highly similar, suggesting that the simple contrast between learning from a single variable speaker and three variable speakers has relatively little impact on behavior. There were two dominant strategies apparent in word learning: regularizing (when participants used one of the labels most of the time) and probability matching (when participants approximated the input frequency, 66%). Participants in the One Speaker condition regularized the Majority Label somewhat more often than in the Within Speakers condition: 27% of all labels in the One Speaker condition were regular, meaning for 27% of the objects, they produced the Majority Label more than 90% of the time (Fig. 2, blue bars on right) as opposed to 17% in the Within Speakers condition (Fig. 2, red bars). These rates of regularization in adult learners are in line with those observed in other unpredictable variation learning experiments with adult learners (around 25% of adults regularize, e.g. in Hudson Kam & Newport, 2005; Smith & Wonnacott, 2010). In both conditions there was a peak around 50-60% Majority Label use, which could reflect slightly inaccurate probability matching behavior or switching between the two labels randomly.

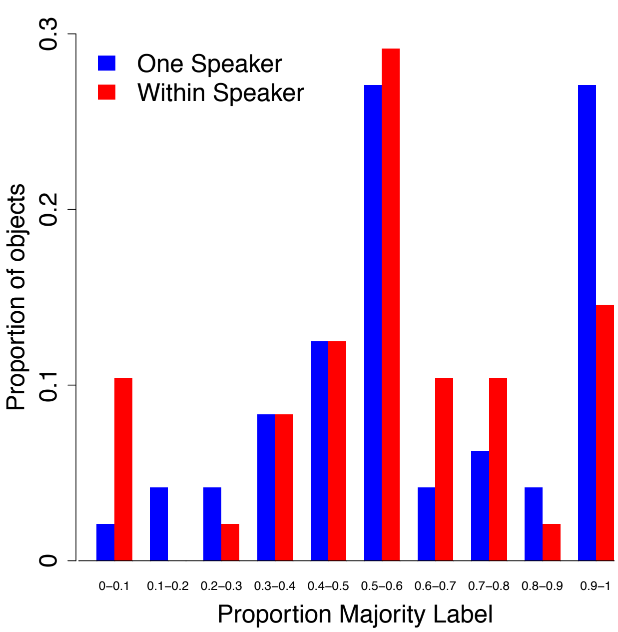

Figure 2: Proportion of objects (y axis) labeled with a given proportion of the Majority Label (x axis, bins) in the One Speaker (blue bars) and Within Speakers (red bars) conditions. Intervals include lower boundary but not upper boundary, except for rightmost bin.

For statistical analysis, multilevel models with random effects were employed: classified responses were fitted to a binomial distribution, implemented in the R programming environment version 3.0.1 using the lme4 package (version 0.99999-2) (Bates et al., 2013). Condition (One Speaker vs Within Speakers) and Testing Block (1-36) were considered as possible fixed effects, with One Speaker set as the baseline condition for comparison. We included Participant, Object, Majority Label and Minority Label as random effects, with random intercepts for each and a by-Participant random slope for Testing Block when Testing Block was included as a fixed effect.

We compared various versions of the model, featuring fixed effects for Condition, Test Block or both, to the equivalent null model featuring the intercept and the random effect structure only (c.f. Mundry, 2011). In no case did the models with fixed effects significantly improve model fit. Models featuring either a fixed effect of Condition or fixed effects of Condition and Test Block were only marginally better that their null equivalents (Fixed effect for Condition only: $\chi2(1)=3.265$, $p=.071$; Fixed effect for Condition and Trial Number: $\chi2(3)=6.965$, $p=.073$), suggesting that there was little reliable difference in propensity to use the Majority Label between conditions. Furthermore, the models featuring fixed effects yielded at best marginally significant differences between the Single and Within Speakers conditions (Condition-only model: $\beta=-0.747$, $SE=0.410$, $p=.066$; Condition plus Training Block model: $\beta=-0.706$, $SE=0.362$, $p=.051$). Consequently, we conclude that there is no reliable difference between conditions, although there may be a slight tendency for more variable performance in the Within Speakers condition, a point we return to in the Discussion.

Experiment 2

In Experiment 2 we ran a further 16 participants through the experiment in the *Between Speakers* condition: as in the Within Speakers condition in Experiment 1, participants received input from all three speakers; however, rather than each speaker exhibiting variability, the variability in labeling was introduced across speakers. Each speaker labeled each object 3 times during each block of training, but each speaker used only one label for each object, and used this label consistently across the entire training set. For each object, two speakers (e.g. A and B) used the Majority Label, and the other speaker (eg. C) used the Minority label. For every object, the identity of the Minority Speaker was different: i.e. each speaker was in the majority in their choice of label on two objects and in the minority on one object.

Results of Experiment 2

We compared the participants in the Between Speakers condition to participants from the Within Speakers condition from Experiment 1. Regularization on the Majority Label was 42% in the Between Speakers condition (Fig. 3, green bars), in contrast with 17% in the Within Speakers condition (Fig 3, red bars). Participants in the Between Speakers condition were more than 5 times as likely to use the Majority Label than the Minority Label, but only twice as likely in the Within Speakers condition. We observed some probability matching behavior in the Between Speakers condition, but for less than 15% of the labels, although matching of the input frequency was more accurate in the Between Speakers condition – this may be due to the fact that the ratio of the Majority to Minority label was tied neatly to the number of speakers in the Between Speakers condition, which may have facilitated accurate learning of the statistics of the input data.

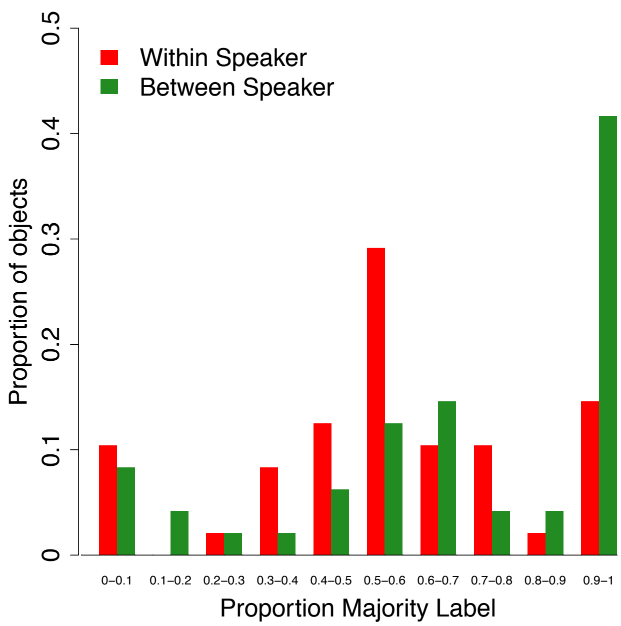

Figure 3: Proportion of objects (y axis) labeled with a given proportion of the Majority Label (x axis, bins) in the Within Speakers (red bars) and Between Speakers (green bars) conditions. Intervals include lower boundary but not upper boundary, except for rightmost bin.

The same statistical techniques were used as in Experiment 1: logit regression with Condition (Within Speakers [baseline] vs Between Speakers) as a fixed effect, with random intercepts for Participant, Object, Majority Label and Minority Label. The model with a fixed effect of Condition did significantly better than the equivalent null model ($\chi2(1)=4.918$, $p=.027$; we also considered a model with a fixed effect of Training Block but this did not improve fit over the simpler model with only a fixed effect of Condition, nor over the equivalent null model).

The estimated intercept ($\beta=0.106$, $SE=0.419$, $p=.8$) suggests that the two labels were used in equal proportions by participants in the Within Speakers condition (in fact there was hardly any difference between the proportion of regularized Majority and Minority Labels), whereas in the Between Speakers condition the Majority Label was used overwhelmingly, for nearly 50% of all objects. Our analysis showed a significant effect of Condition ($\beta=1.011$, $SE=0.445$, $p=.023$): participants in the Between Speakers condition were significantly more likely to produce the Majority Label than participants in the Within Speakers condition.

Discussion

Participants' responses to unpredictable variation are influenced by how that variation is associated with social information in their input. We found no significant difference between learning from one or multiple teachers. However, Experiment 2 revealed a substantial difference between the Within Speakers and the Between Speakers conditions: participants were significantly more likely to use the Majority Label when variation was distributed across speakers.

There are at least three non mutually exclusive possible explanations for this behavior. Firstly, this behavior is consistent with conformist copying of labels, suggesting that this domain-general social learning strategy also operates during word learning, and leads (in the right social scenario) to the elimination of linguistic variation. However, it is also possible that learners track the evidence for within-speaker variability and adjusting their own variability accordingly: whereas the Within Speakers condition provides good evidence that speakers are truly variable, the Between Speakers condition provides evidence that individual speakers are not variable. Participants in this condition might therefore be more likely to be consistent and regular, and might be more likely to choose the Majority Label simply because of its higher frequency in their input. This explanation is consistent with accounts that children are socially sensitive to their interlocutors' preferences when learning novel linguistic input (Chang et al., 2009). If this explanation were correct, however, we might also expect to see differences between the One and Within Speakers conditions in Experiment 1: the Within Speakers condition provides better evidence (across multiple teachers) of within-speaker variability. If present in Experiment 1, this effect is, however, extremely weak compared to the effect observed in Experiment 2.

Finally, there may also be memory-based accounts for the difference between the Within and Between Speakers conditions. Hudson Kam and Chang (2009) found that adults probability matched more when retrieval was facilitated (e.g. by providing vocabulary on flash cards), therefore regularization was a result of difficulty recalling the words rather than encoding them. In contrast, Perfors (2012) showed that increased memory load does not result in increased regularization. In the Between Speakers condition, each Minority label was only produced by a single speaker, and therefore may have been harder to learn than the Majority labels which were produced by two speakers. It has been shown that learning from multiple speakers aids phonological processing and learning (Rost & McMurray, 2009). In addition, varying the Minority speaker for each object also substantially increased the cognitive demands of the task.

Further experiments are required to tease apart these various possible explanations: conformist learning, learners inferring the within-speaker variability of their language, and lower-level memory demands all might play a role in our results. It is also possible that paradigms in which learners receive input from multiple teachers might lead to rather different regularization dynamics over iterated learning (Reali & Griffiths, 2009, Smith & Wonnacott, 2010): using a more complex iterated learning paradigm in which people learn from multiple teachers to study the process of this would be a useful addition to this literature.

Conclusions

We have shown that social information plays an important role in how people respond to linguistic variation: in particular, learners who encounter variation distributed between, but not within, speakers act in a conformist manner, preferentially adopting the variant used by the majority of their teachers. There are several possible explanations for our results, including social conformity, which is an important learning mechanism in other socially acquired behaviors. Alternatively, learners may track the variability in their teachers' behavior, or lack thereof, and adjust their own behavior accordingly. Exploring these possibilities further will shed light on social learning mechanisms employed in language learning, and how and when these biases act to maintain or eliminate linguistic variation.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version*, *1*(4).

Chang, F., Kobayashi, T., & Amano, S. (2009). Social factors in the acquisition of a new word order. *First Language*, 29, 427-445.

Givón, T. (1985). Function, structure, and language acquisition. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 2, pp. 1005–1028). Hillsdale, NJ: Lawrence Erlbaum.

Hudson Kam, C., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 815–821.

Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, 59, 30-66.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1, 151–195.

Johnson, J., Shenkman, K.,Newport, E. L., & Medin, D. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language, 35,* 335–352.

Labov, W. (1963). The social motivation of a sound change. *Word*, 19, 273–309.

Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W. J. E. & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B*, 279, 653-662.

Mundry, R. (2011). Issues in information theory-based statistical inference – a commentary from a frequentist's perspective. *Behavioral Ecology and Sociobiology, 65*(1), 57-68.

Pike, T. W. & Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biology Letters*, 6, 466-468.

Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, 111, 317–328.

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

Rost, G. C. & McMurray, B. (2000). Speaker variability augments phonological processing in early word learning. *Developmental Science*, 12, 339-349.

Sankoff, G. (1979). The genesis of a language. In K. C. Hill (Ed.), *The genesis of language*, Ann Arbor, MI: Karoma Publishers.

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from impoverished input. *Cognitive Psychology*, *49*, 370–407.

Wonnacott, E., & Newport, E. L. (2005). Novelty and regularization: The effect of novel instances on rule formation. In A. Brugos, M. R. Clark- Cotton, & S. Ha (Eds.), *BUCLD 29: Proceedings of the 29th annual boston university conference on language development.* Somerville, MA: Cascadilla Press.

Van de Waal, E., Borgeaud, C. & Whiten, A. (2013). Potent social learning and conformity shape a wild primate's foraging decisions. *Science*, 340, 483-485.