Language Acquisition

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hlac20

Social biases can lead to less communicatively efficient languages

Masha Fedzechkina, Lucy Hall Hartley & Gareth Roberts

To cite this article: Masha Fedzechkina, Lucy Hall Hartley & Gareth Roberts (2022): Social biases can lead to less communicatively efficient languages, Language Acquisition, DOI: 10.1080/10489223.2022.2057229

To link to this article: https://doi.org/10.1080/10489223.2022.2057229

Published online: 27 Jun 2022.

LANGUAGE ACQUISITION
https://doi.org/10.1080/10489223.2022.2057229

Routledge
Taylor & Francis Group

Social biases can lead to less communicatively efficient languages

Masha Fedzechkina[a], Lucy Hall Hartley[b], and Gareth Roberts[c]

[a]Department of Linguistics, Graduate Interdisciplinary Program in Cognitive Science, Graduate Interdisciplinary Program in Second Language Acquisition and Teaching, University of Arizona; [b]Department of Linguistics, University of Arizona; [c]Department of Linguistics, University of Pennsylvania

ABSTRACT

Language is subject to a variety of pressures. Recent work has documented that many aspects of language structure have properties that appear to be shaped by biases for the efficient communication of semantic meaning. Other work has investigated the role of social pressures, whereby linguistic variants can acquire positive or negative evaluation based on who is perceived to be using them. While the influence of these two sets of biases on language change has been well documented, they have typically been treated separately, in distinct lines of research. We used a miniature language paradigm to test how these biases interact in language change. Specifically, we asked whether pressures to mark social meaning can lead linguistic systems to become less efficient at communicating semantic meaning. We exposed participants to a miniature language with uninformative constituent order and two dialects, one that employed case and one that did not. In the instructions, we socially biased participants toward users of the case dialect, users of the no-case dialect, or neither. Learners biased toward the no-case dialect dropped informative case, thus creating a linguistic system with high message uncertainty. They failed to compensate for this increased message uncertainty even after additional exposure to the novel language. Case was retained in all other conditions. These findings suggest that social biases not only interact with biases for efficient communication in language change but also can lead to linguistic systems that are less efficient at communicating semantic meaning.

ARTICLE HISTORY

Received 1 August 2021
Accepted 16 February 2022

Introduction

An important function of language is to efficiently convey information about events. For this to succeed, utterances must provide low uncertainty about the semantic roles of those involved in the events (i.e., who is doing what to whom). In a sentence like “Asterix provoked Caesar,” English constituent order leaves little uncertainty as to who did the provoking. Other languages might achieve the same goal by different means. Classical Latin, for example, allowed much more constituent order flexibility: In Latin, *Asterix Caesarem provocavit, Caesarem Asterix provocavit*, and *provocavit Caesarem Asterix* all denote the same event. To distinguish who was doing what to whom, classical Latin relied on case markers (morphological elements on nouns and pronouns that indicate their role in the sentence, such as the *-em* ending on *Caesar* in our example). Constituent order and case marking are not the only means of doing this. A variety of grammatical devices are put to the task in different languages, such as agreement (e.g., marking properties of the subject and object on the verb, as in Nahuatl; Launey, 2011), prosody (e.g., in German; Weber, Grice, & Crocker, 2006), or other modifications of the object itself (e.g., “mutating” its initial consonant under certain circumstances in Welsh; Tallerman, 2006).

CONTACT
gareth roberts@ ng upenn edu
Gareth Roberts

While a number of different mechanisms for distinguishing semantic roles may coexist in the same language, no known language makes use of all such mechanisms, and they tend to be distributed in a rather complementary fashion (Van Everbroeck, 2003). Perhaps most strikingly, it has long been observed that there exists a trade-off such that languages with more fixed constituent order (e.g., English, French, or Mandarin) tend to exhibit less case marking, while languages with more flexible constituent order (e.g., Russian, Latin, or Turkish) tend to have more case marking (Koplenig, Meyer, Wolfer, & Mueller-Spitzer, 2017; Levshina, 2021; Sapir, 1921). Recent information-theoretic work has linked this cross-linguistic pattern to the principle of balancing uncertainty against production effort (Jäger, 2007; Kurumada & Jaeger, 2015). Under this principle, languages with fixed constituent order —in which semantic roles can be reliably inferred from constituent order alone—are unlikely to employ redundant case marking, thereby reducing the production effort associated with producing additional morphemes without sacrificing robust message transmission.[1] Languages with flexible constituent order—in which constituent order alone is not sufficiently informative—recruit an additional cue to semantic roles, such as case, to reduce uncertainty about the intended message at the expense of an increase in production effort.

This explanation has received experimental support from studies employing a miniature-language learning paradigm in which participants learn a novel artificial language and then produce sentences in the language to describe events (Fedzechkina & Jaeger, 2020; Fedzechkina, Newport, & Jaeger, 2017; Hall Hartley & Fedzechkina, 2020). The language to which participants are exposed typically affords several options to describe the same event, some of which are consistent with cross-linguistically common patterns, while others are not (Culbertson, Smolensky, & Legendre, 2012; Fedzechkina, Newport, & Jaeger, 2016; Hudson Kam & Newport, 2009; Smith & Wonnacott, 2010). For example, Fedzechkina and Jaeger (2020) trained participants on miniature languages with variable case marking and manipulated both the effort required to produce case markers (operationalized in terms of mouse clicks) and uncertainty about the intended message (by varying constituent order flexibility). They found that learners changed the input to maintain case in the flexible order language and to drop case in the fixed order language only when case production required additional effort compared with a non-case-marked noun. This led languages to better balance uncertainty against production effort and made them more consistent with the cross-linguistically observed trade-off between case and constituent order flexibility.

However, conveying semantic roles is not the only function of the grammatical devices mentioned so far: Languages recruit the same devices to carry other information as well. For example, both constituent order (as in, e.g., Hungarian; Puskás, 2000) and case markers (as in, e.g., Japanese; Hasegawa, 2011) can be used to mark information-structural elements, such as topic and focus.[2] The same devices can also be recruited to convey social meaning about the language users and their relationships and attitudes. Indeed, whenever there is linguistic variation, grammatical devices can acquire social significance by becoming associated with the speakers who most use them, or with stereotypical characteristics of those speakers (Eckert, 2008). For instance, features of Southern US speech (such as the word *y'all*) might acquire positive associations of warmth but negative associations of lack of education, both widespread stereotypes of the American South

[1] Here we are assuming that case marking involves extra morphemes. This is not strictly speaking necessary; case can also be marked by phonological alternations on the root (as, for instance, in Irish). However, this is a less common pattern, and the alternations involved often arise historically under the influence of case-marking morphemes that were subsequently lost. There is also a cost associated with maintaining complex case paradigms, however marked (Ackerman & Malouf, 2013).

[2] The Japanese morphemes in question are more typically referred to as discourse particles, rather than case markers, owing to the nature of what they mark; our point here is that similar post-nominal morphemes in Japanese are used to mark both case and information structure.

(Preston, 1998, 1999). Such social meaning can attach to practically any part of language, including the grammatical devices for conveying semantic meaning described above. For example, the English form *whom*, which was originally a case-marked form of *who*, has been co-opted to primarily mark social meaning (e.g.,

education or pretension) in modern English (Lasnik & Sobin, 2000). This social meaning marking has consequences for the use and propagation of the forms involved (Eckert, 2008; Sneller & Roberts, 2018), causing speakers to adopt or avoid variants depending on the social effect they want to achieve.

In other words, linguistic units are used to convey multiple kinds of meaning simultaneously, including but not limited to information about events or states in the world (which we will call *semantic* meaning) and social information about the speaker (which we will call *social* meaning). This has the consequence that any given instance of linguistic communication is likely to involve a somewhat complex juggling of resources for the purpose of efficiently achieving the language user's communicative goals. In other words, language users must make production choices that ensure their intended audience successfully infers both the intended semantic and the intended social meaning. Importantly, the successful communication of social meaning is orthogonal to the successful communication of semantic meaning. While they need not be at odds, and may coincide under some circumstances, they may under other circumstances push language users in different directions, such that satisfying the goal of communicating one meaning can lead to potential uncertainty about the other (Labov, 2001, pp. 3–6). Socially driven avoidance of the word *y'all*—whose use allows number distinctions to be made in English second-person pronouns—is a good example of this. The avoidance of *y'all* (and similar forms such as *yinz* and *yous*) is highly typical of socially prestigious registers, often leaving it unclear whether *you* refers to one person or more (Preston, 2015). The form *whom*, by contrast, is widely promoted in the same registers, even though constituent order very rarely leaves the grammatical or semantic role of *who* uncertain.

But how precisely do such social biases and biases for efficient communication of semantic meaning interact in shaping language change? We used a miniature language learning paradigm to investigate this. Our central question is to what extent pressures for the efficient communication of social meaning can reduce or even reverse the effects of pressures for the efficient communication of semantic meaning. Roberts and Fedzechkina (2018) made a first step in exploring this question. In their experiment, which had an iterated-learning design (in which *generations* of participants learn a language based each time on the output of the previous generation; cf. Kirby, Griffiths, & Smith, 2014), participants learned a miniature "alien" language with two dialects. Both dialects shared 100% consistent subject-object-verb (SOV) constituent order but differed with regard to case marking: While one dialect had none, the other dialect had 100% consistent case marking on the object. Thus, case marking in the language overall was redundant—as the semantic meaning could be reliably inferred from constituent order alone—and socially conditioned. In the instructions, Roberts and Fedzechkina biased participants to feel positively inclined toward speakers of one of the two dialects, toward speakers of both dialects, or against speakers of the case dialect. They found that the redundant case marker disappeared rapidly in all conditions, but its loss was considerably slower when participants were biased toward users of the case-marking dialect.

Roberts and Fedzechkina's (2018) study established a paradigm for investigating how social and other biases might interact in language change. However, their study focused on the retention of *redundant* case marking: Whether it disappeared or was retained, the communication of the intended semantic meaning was barely affected, as it could be reliably inferred based on constituent order alone.[3] The more interesting question, perhaps, is what happens when *informative* case marking, important for reducing uncertainty about semantic meaning (such as in a language with uninformative constituent order), acquires social meaning. Would a social bias in favor of dropping case lead to the loss of case marking in such circumstances, thus creating an increased uncertainty

[3] This is not to say that redundant morphology carries no information; on the contrary, it plays an important role in combating noise (Frank & Jaeger, 2008; Stevens & Roberts, 2019). For the rather constrained communicative contexts in our setup, however, noise levels are low enough for a single grammatical cue to be generally sufficient.

in the linguistic system about the intended semantic meaning? If so, would speakers develop alternative strategies to reduce the increased uncertainty associated with case drop (such as by fixing constituent order)?

Here, we investigated this question in two experiments by exposing participants to a language with flexible (i.e., uninformative) constituent order and dialectal variation in the presence or absence of informative case marking. Following Roberts and Fedzechkina (2018), we manipulated social biases in the first experiment as a between-participant variable, biasing different groups of participants to the dialect with case marking, to the dialect with no case marking, or to neither of the two dialects. After exposure, participants produced novel sentences in the language they had learned. We measured whether participants' own use of case marking was affected by the social bias. In the second experiment we further probed whether these preferences changed after more extensive experience with the novel language and whether learners introduced changes into the linguistic systems to compensate for the increased uncertainty about semantic meaning (which we will term *message uncertainty*) associated with case loss.

Experiment 1

Participants

Recruitment and execution of this study was approved by the Human Subjects Protection Program at the University of Arizona. Participants were recruited through Prolific, a crowd-sourcing platform. Participants were prescreened to be (self-reported) monolingual speakers of English with no known language disorders who had at least 95% past approval on Prolific. The experiment was administered via FindingFive, a platform for the design and administration of behavioral experiments online (Finding Five Corporation, 2019).

Each participant was exposed to only one condition, which lasted approximately 50 minutes, and received $7 for participation. In line with prior work (Fedzechkina, Chu, & Jaeger, 2018; Fedzechkina, Jaeger, & Newport, 2012), participant recruitment continued until the number of participants who had successfully learned the miniature language reached 20 in each condition. Successful learning was defined exactly as by Fedzechkina et al. (2018), which reduced our degrees of freedom in deciding when to stop recruitment (see Scoring and Exclusions section below for details). The final sample submitted for analysis included 60 participants (out of 96 participants who completed the experiment).[4]

Miniature input language

Participants were informed that they would learn an "alien" language by watching short videos accompanied by sentences that described them in the novel language. The language contained four novel nouns (*barsa, dokla, koofta, pilka*) that corresponded to humanoid referents (CHEF, MOUNTIE, REFEREE, BANDIT), two novel verbs (*kyse, tegut*) that corresponded to transitive actions (KICK and HUG), and a case suffix *-dak* that (if present) attached to the object. The novel words (all phonotactically legal in English) were generated separately using the Apple speech synthesizer (voice "Alex") and concatenated into sentences using a Praat script, thus ensuring that no prosodic cues to sentence meaning were present.

Participants were instructed that there were two species of aliens (each with its own color—orange or blue) that spoke slightly different dialects. Both dialects had flexible constituent order:

[4] Post-hoc power analysis conducted to determine participant numbers for Experiment 2 (see Experiment 2 *Participants* section) further confirmed that this final sample had appropriate power to detect all effects considered in Experiment 1 (80% as recommended by Brysbaert & Stevens, 2018).

Subject- object-verb (SOV) and object-subject-verb (OSV) occurred equally frequently in each of them. The dialects differed, however, in how they employed case marking (Figure 1). The *case dialect* had a case marker (the suffix *-dak*) on every object-noun. The *no-case dialect* had no case marking on any noun. Thus, case marking in the language *was* dependent on dialect while constituent order *was not*. Overall, this meant that 50% of the sentences that participants were exposed to had SOV constituent order,

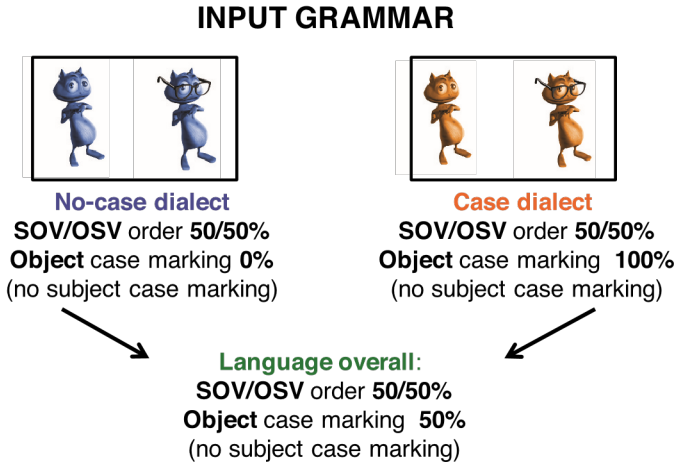


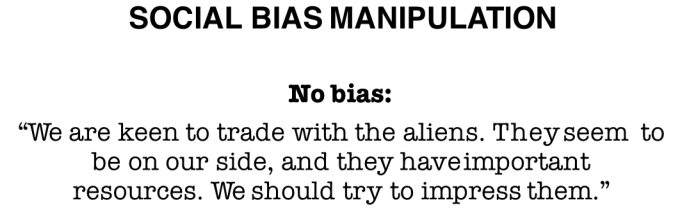
Figure 1. Schematic illustration of the miniature language grammar used in both experiments. Participants were exposed to two alien language dialects (indicated by alien color—blue or orange). Both dialects had flexible constituent order; one of the dialects employed case; the other dialect did not. Overall, learners were exposed to a flexible order language with variable case marking.

while the other 50% had OSV constituent order; and 50% of the sentences for each constituent order (and by design, 50% of sentences overall) had object case marking. Such a language makes it impossible to tell who is doing what to whom based on sentence constituent order alone. Object case marking, when present, eliminates this uncertainty. During training, every video was accompanied by a picture of the alien informant to indicate which dialect the utterance came from.

All verbs occurred equally frequently with both constituent orders and all nouns occurred equally frequently as subject and object with each verb. To avoid unintentional associations between the novel labels and meanings, their assignment to meanings was rotated across two lists. Scenes were accompanied by both auditory and written descriptions (these were necessary to familiarize learners with the spelling of the alien words, as they produced languages by typing).

Social bias manipulation

Participants were randomly assigned to one of three conditions differing only in the instructions given to participants at the start of the experiment, which encouraged them to feel positively inclined toward one or both alien species (Figure 2). Participants in all conditions were explicitly told that there were two groups of aliens who spoke two slightly different dialects and could be distinguished by color (one group was blue and the other was orange). In the instructions for the *no-bias* condition, participants



Dialect bias (case or no case):

"We are especially keen to trade with the blue aliens. They seem to be on our side, and they have important resources. We should try to impress these blue aliens in particular."

Figure 2. Instructions used to bias learners to feel positively toward either the alien speakers of a particular dialect (case or no-case) or the alien speakers overall. The key parts are underlined (underlining was not shown to participants).

were encouraged to feel positively about both groups of aliens. In the other conditions, they were encouraged to feel positively inclined toward only one of the two groups of aliens—either the speakers of the case dialect (in the *bias-for-case* condition) or the speakers of the no-case dialect (in the *bias-for- no-case* condition).

Throughout the instructions, including in the bias text, the aliens were referred to by color only; no reference was made to any feature of the aliens' language. During the grammar learning part of the experiment (i.e., during sentence exposure and sentence comprehension), participants saw short videos of humanoid characters performing simple transitive actions, along with a picture of one of the aliens on the bottom left of the video. The picture of the alien was accompanied by a speech bubble to indicate that the sentence accompanying the video was produced by this alien species. No aliens were presented during the sentence production test (see Figure 3 and *Procedure* section). The

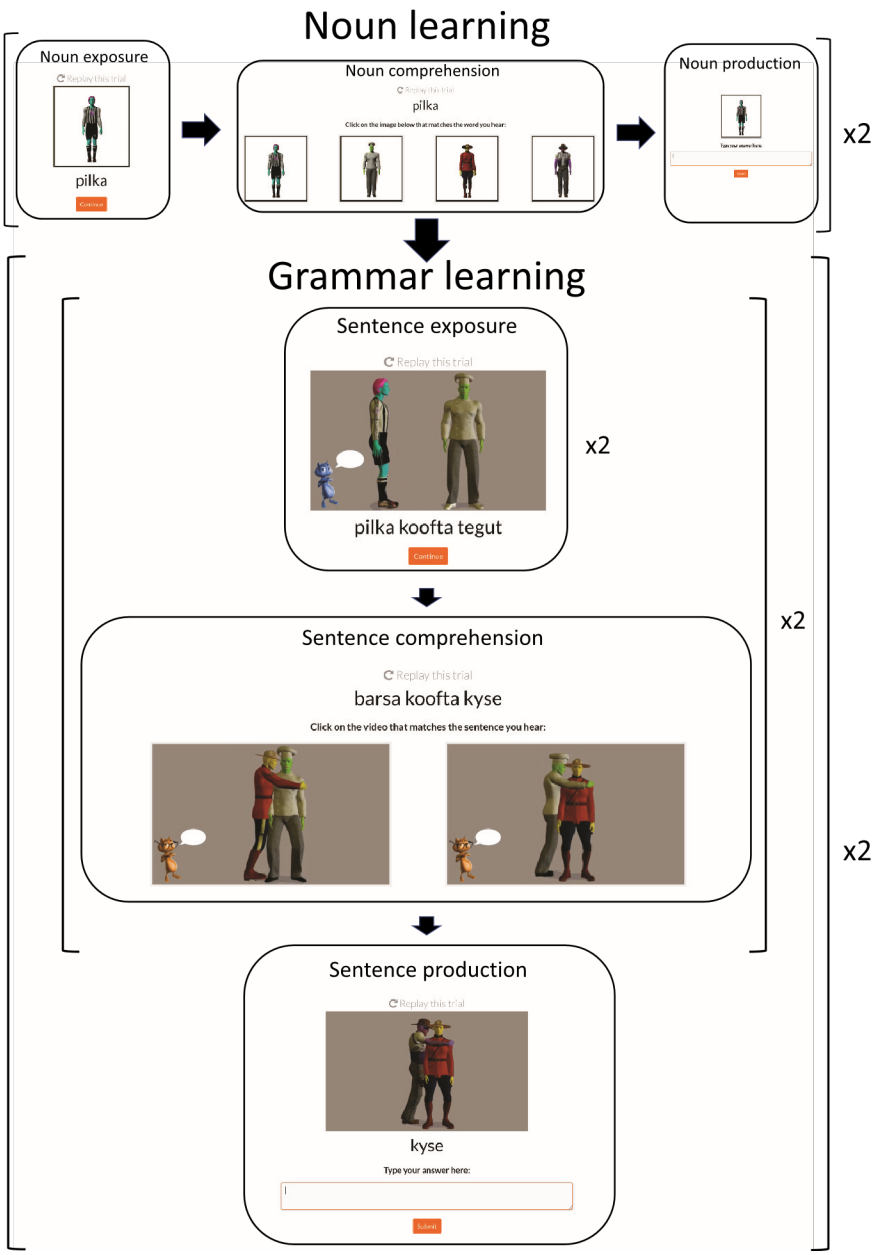


Figure 3. Full experiment procedure with sample screenshots (Exp. 1). Noun exposure and noun comprehension consisted of eight trials each, while noun production consisted of four trials. Each block in grammar learning consisted of 16 trials. Alien speakers can be seen on the bottom left of the videos in sentence exposure and sentence comprehension, but not sentence production. Exp. 2 followed an identical procedure on days 1 and 2.

manipulation was identical to that of Roberts and Fedzechkina (2018) with one exception. To ensure that their instructions were indeed inducing a social bias toward a group of aliens rather than simply drawing attention to their dialect, Roberts and Fedzechkina included an additional condition that biased learners *against* the speakers of the case dialect, while explicitly mentioning this group (and thus directing attention toward it). This control condition was omitted here because, in Roberts and Fedzechkina's study, it had precisely the same effect as the bias in favor of users of the no-case dialect, suggesting that the instructions were indeed inducing a social bias rather than merely directing attention to a dialect. We discuss this in more detail in the General Discussion section.

Procedure

At the beginning of the experiment, participants were informed that they would be learning a novel alien language by watching short videos describing simple events accompanied by descriptions of them in the novel language. Participants were also informed that the language had two different dialects spoken by different species of aliens. Depending on the condition, they were also encouraged at this point to feel positively about one or both alien species (see *Social Bias Manipulation* section). They were not, however, provided with any information about the grammar of the language or the linguistic differences between the dialects.

The experiment was organized into two phases— *noun learning* and *grammar learning* (Figure 3).

Noun learning. The experiment began by teaching participants the names for the humanoid characters involved in the different scenes. This noun learning phase consisted of three blocks of trials. The first block was *noun exposure*. In this block, participants viewed a picture for each of the characters one at a time, accompanied by a label in the novel language (presented both auditorily and in writing). Each of the four characters was seen twice, resulting in eight trials in total. After noun exposure, participants performed a *noun comprehension test*. Participants heard a novel label and were asked to choose the corresponding picture out of an array of all four characters. After each trial of the noun comprehension block (eight in total), they received feedback on their accuracy. Finally, after the noun comprehension block, participants completed a *noun production test*, during which they were asked to provide the name for each of the characters once. As with noun comprehension, participants received feedback on their accuracy on each trial. Participants completed the noun learning phase twice before moving to the grammar learning phase.

Grammar learning. Like noun learning, grammar learning consisted of three blocks of trials. First, in *sentence exposure*, participants learned the grammar of the language by watching short videos depicting simple transitive events performed by two humanoid characters (e.g., a chef hugging a referee). Each video was accompanied by a picture of an alien informant (from either the blue or the orange species) to indicate which dialect the accompanying sentence came from. The sentence exposure block was repeated twice in a row, with 16 trials each time. After the sentence exposure blocks, participants completed a *sentence comprehension test* consisting of 16 trials, in each of which they heard and saw a sentence accompanied by two videos. Each of the two videos involved the same two characters in reversed semantic roles (that is, the actor in one video was the patient in the other). Participants were asked to click on the video that matched the sentence they heard. As in sentence exposure, the videos included a picture of the alien informant to indicate the dialect used in the sentence. No feedback was provided on sentence comprehension trials.

Upon completing the sentence comprehension test, participants proceeded to a *sentence production test* (the critical test in our experiment). This block also consisted of 16 trials, in each of which participants were asked to type sentences in the alien language to describe previously unseen videos. To make this task easier, they were prompted with the alien-language verb both auditorily and in writing. Sentence production videos did not contain a picture of the alien informant, so participants were free to choose the dialect they wanted to use. No feedback was provided on sentence production.

Participants completed the grammar learning phase twice; each time, it consisted of the same blocks in the same order (Figure 3). This means that each participant experienced eight blocks of sentence exposure, four sentence comprehension blocks, and two sentence production blocks in total. Throughout the experiment, participants could replay the videos and the sentences that went with them as many times as they liked.

Results

Before we turn to the main question of our study—whether learners introduced changes into the input case-marking distribution as a result of social biases—we describe our data scoring method, our participant exclusion criteria, and the accuracy of acquisition.

Scoring and exclusions

We recorded learners' accuracy on sentence comprehension and production trials, along with learners' case marking and constituent order preferences on sentence production tests. For sentence comprehension trials, we assessed participants' accuracy on case-marked (i.e., unambiguous) trials. Since sentence constituent order was uninformative about semantic meaning, accuracy on case-marked trials indicated how well participants learned the meaning of case marking. Following Fedzechkina et al. (2017) and Fedzechkina and Jaeger (2020), participants who failed to reach 70% accuracy on the final comprehension test were removed from the analysis. This included seven participants in the bias- for-case condition, nine participants in the bias-for-no-case condition, and 17 participants in the no- bias condition.

All production trials (noun and sentence) were automatically annotated for accuracy using a custom Python script. Lexical items were considered correctly labeled in the alien language if they were within a Levenshtein distance of two of the target (i.e., we allowed at most two character insertions, deletions, or substitutions in a word). For example, "togla" would be considered a correct label for "dokla," but "togli" would not. For each sentence produced by participants, we recorded the number of lexical mistakes (i.e., lexical items within a Levenshtein distance greater than two of the target) they made. If there was more than one such mistake, the sentence was scored as "uncodable" (since we could not reliably determine the constituent order intended by the participant) and removed from further analysis. Participants with at least 50% uncodable productions (three participants total, two in the bias-for-case and one in the bias-for-no-case condition) were excluded from further analyses. This left a total of 60 participants for analysis, 20 in each social bias condition.

For every codable sentence, we annotated which constituent order was used, whether the case marker was present, and what constituent the case marker was attached to. All sentences containing a grammatical mistake (i.e., using a constituent order other than SOV and OSV or using a case marker on a constituent other than the object) were excluded from all analyses.

Accuracy of acquisition

For participants included in the analysis, the accuracy of both lexical and grammar acquisition was high. In the final sentence production test, participants made grammatical errors on less than 3% of sentences (see Table 1) and lexical errors on less than 1% of sentences in each condition. Similarly, comprehension accuracy on unambiguous (i.e., case marked) trials was high on the final comprehension test—above 95% in each condition (see Table 2). This suggests that, despite its difficulty (as indicated by the high exclusion rates), the task overall was feasible for our participants.

We now turn to our two main questions: Did learners drop case marking as a result of a social bias, despite it being informative in the input, and if so, did they adopt other strategies to make up for the increased message uncertainty?

Table 1. Grammatical errors in production in Exp. 1

		Case errors		Constituent order errors	
Test	Bias condition	Mean	95% CI	Mean	95% CI

Test 1	Case	1.88%	0.31-4.06%	5.00%	0-15.00%
	No-bias	0.94%	0-1.88%	4.38%	0-13.13%
	No-case	3.12%	0.31-6.88%	0.31%	0-0.94%
Test 2	Case	0.31%	0-0.94%	0.63%	0-1.56%
	No-bias	1.25%	0-2.81%	0%	0-0%
	No-case	2.18%	0-5.63%	0%	0-0%

Table 2. Comprehension accuracy in Exp. 1

		Accuracy	
Block	Bias condition	Mean	95% CI
Block 1	Case	70.00%	61.88-78.13%
	No-bias	73.75%	61.25-85.00%
	No-case	77.50%	63.75-88.75%
Block 2	Case	81.25%	67.25-93.13%
	No-bias	89.38%	79.36-97.50%
	No-case	92.50%	83.75-98.75%
Block 3	Case	93.13%	89.38-96.25%
	No-bias	99.38%	98.13-100%
	No-case	91.25%	83.13-97.50%
Block 4	Case	95.63%	92.48-98.14%
	No-bias	98.75%	96.88-100%
	No-case	98.75%	96.88-100%

Case use in production

To address our first question—whether learners dropped informative case marking as a result of a social bias—we conducted two analyses: We compared case use between social bias conditions and we also compared case use with the input. These analyses present a complementary picture of learners' preferences, as learners might have different preferences across conditions while at the same time not deviating significantly from the input they receive (or the other way around).

We used mixed effects logistic regression to predict the presence of case marking from the social bias condition (sliding difference coded:[5] no-bias vs. bias-for-case; bias-for-no-case vs. no-bias condition), production test block (sum-coded, 2 vs. 1), and their interaction. The model contained the fullest converging random effects structure (random intercepts for participant and item—defined as object-noun—and by-participant random slope for production test block). It revealed a main effect of production cantly more test case block marking on case on the use, second $\hat{\beta}$ à 0*:*test, 86, *z*after à 2*:*9they , *p* àhad 0*:*003become , meaning more that proficient learners with used the signifinovel language.[6] This is consistent with prior work using similar artificial languages (Fedzechkina & Jaeger, 2020; Fedzechkina et al., 2017). There was no significant difference in case

[5] Sliding difference coding compares the mean of the dependent variable for one level of the categorical variable to the mean of the dependent variable for the prior adjacent level. This coding scheme is appropriate for this analysis since the conditions in our experiment are ordered in terms of the expected likelihood of case use: bias-for-case > no-bias > bias-for-no-case. [6]See Appendix A for full results of this model and for all other models reported.

use between the bias-for- case and the no-bias conditions (*β*^à dialect 0*:*3, *z*was à not 0*:*61strong , *p* à 0enough *:*54; see to Figure lead learners 4), suggesting to deviate that from the social bias toward speakers of the case the input beyond the baseline (i.e., the no-bias condition). On the other hand, learners in the bias-for- no-case condition used significantly less case compared with the no-bias condition (*β*^ à 1*:*65,

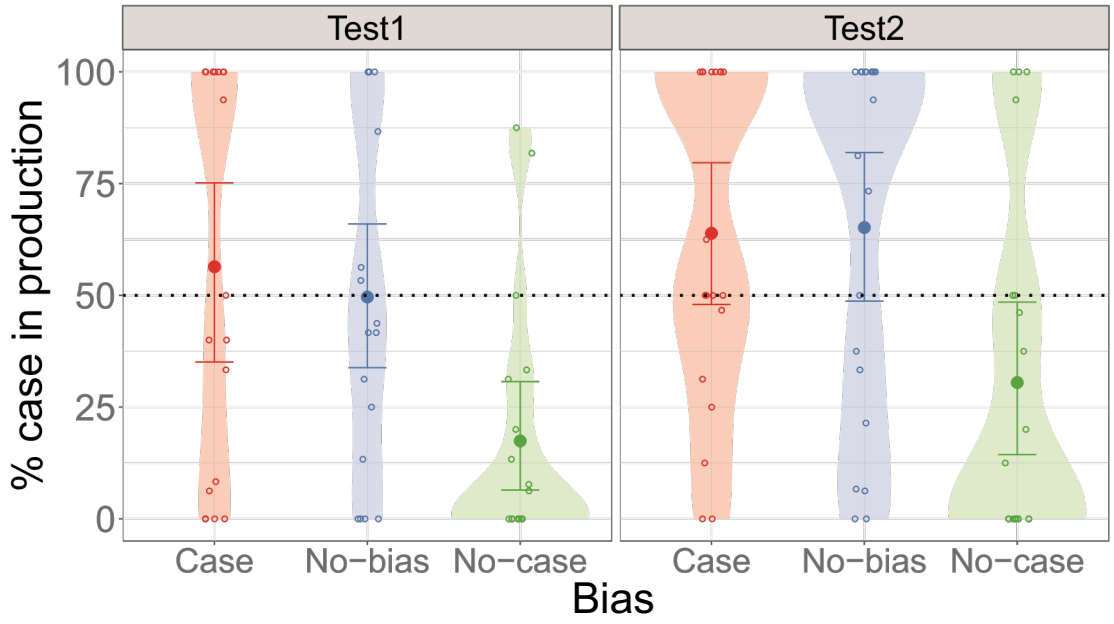


Figure 4. Case use in production by social bias condition in Exp. 1. The dashed line represents the input proportion (same across social bias conditions). The large dots represent condition means. The small dots represent individual participant means. The error bars represent bootstrapped 95% confidence intervals.

z àa decrease 3:12, *p*in àcase 0:002)use , suggesting compared that with a the social other bias groups. toward There the speakers were no of other the no-case-dialect significant effects resulted in the
in model (smallest $p >$
0:4).

To further understand how learners used case marking in the three bias conditions, we compared learners' case use with the input on the second sentence production test. We used mixed effects logistic regression to predict the presence of case marking from social bias condition (treatment coded) with the fullest converging random effects structure (random intercepts for participant and item). The intercept of this model captures whether the social bias condition coded as the reference level significantly differs from 0.5, our input proportion of case. We ran this model three times, with each social bias condition coded as the reference level. These analyses revealed that learners in the bias-for-case condition did not differ from the input

proportion learners significantly in the (63more %; bias-for-no-case $\hat{\beta}$ case à 2:compared 70, *z* àcondition 1:92to , the *p* àproduced input 0:053)(65. Learners case %; β^significantly 2in :77the , *z* àno-bias below 1:97, the *p*condition àinput 0:048)level . produced Notably, (30%; à

β^àThese 3*:*26data , *z* àthus 2*:*22suggest , *p* à 0that *:*026).learners produced less case marking when they were socially biased toward speakers of the no-case dialect. However, since case marking was an informative cue in the language (as the semantic meaning could not be reliably determined based on constituent order alone), producing less case marking in the bias-for-no-case condition could result in increased message uncertainty compared to the other conditions. Note, however, that the miniature language input allowed learners several pathways to avoid the increased uncertainty in the bias-for-no-case condition (such as by fixing constituent order, thus making it informative). We discuss next whether learners took advantage of these possibilities and avoided increased message uncertainty in the bias-for-no-case condition compared to the other conditions.

Message uncertainty in production

Our miniature language afforded two main pathways for language change that could reduce message uncertainty caused by decreased case use: fixing constituent order or conditioning case use on constituent order (or some combination of the two). Prior work has shown that learners in these types of experiments vary significantly in the strategies they employ (Fedzechkina et al., 2017). Thus, to capture the amount of uncertainty about the intended meaning independently of the particular strategies participants employed, we calculated message entropy in each participant's output as the entropy of constituent order in non-case-marked sentences weighted by the proportion of non-case- marked sentences.

Given the input grammar, case-marked sentences contain no uncertainty about the intended meaning. Thus, minimal message entropy (0 bits) is achieved by all systems that have no constituent order variation (regardless of the presence of case marking), by systems that have consistent case marking (regardless of constituent order variation), or by systems that maintain constituent order flexibility while consistently using case marking with only one constituent order variant. Maximal message entropy of 1 bit is achieved in a system that has two constituent orders used with equal frequency and no case marking. The remaining possible systems, given our input, fall somewhere in between. Consider the miniature input language that participants were exposed to (Figure 1). Fifty percent of input sentences contained case marking, resulting in message entropy of 0 bits, and 50% of input sentences contained no case marking while maintaining maximal constituent order flexibility, resulting in message entropy of 1 bit. Thus, the overall message entropy of the input was 0.5 bits (constituent order entropy in non-case-marked sentences of 1 * proportion of non-case-marked sentences of 0.5 = 0.5; the proportion of case-marked sentences is not included in the calculation, as they contribute 0 bits to the overall system entropy).

We used linear regression to predict conditional entropy in production from social bias condition, production test block, and their interactions. The variables were coded in the same way as in the Case Use in *Production* section. The residuals from our model were not normally distributed, so we transformed our conditional entropy data using the R package bestNormalize (Peterson, 2021), which determined that an exponential transformation was the best fit. Our linear regression with the transformed data revealed that learners in the no-bias condition produced linguistic systems that did not significantly differ in conditional entropy from those produced by learners in the bias-for-case condition (*β*^à 0*:*05linguistic , *t* à 0*:*73systems , *p* à 0*:*with 47; see significantly Figure 5). However, higher conditional learners in entropy the bias-for-no-case (i.e., higher

condition produced

uncertainty) than learners in the no-bias condition ($\hat{\beta}$ à 0*:*30, *t* à 4*:*34, $p < 0.0001$). There were no other significant effects (smallest $p > 0.3$).

We further compared the message uncertainty in the linguistic systems produced by learners to the input message uncertainty of 0.5 bits. Looking at the second test block only, we ran a linear

regression that predicted conditional entropy (again with the exponential transformation) from social bias condition (treatment coded) with an offset of 0.5 corresponding to the input conditional entropy. The intercept of this model captures whether the reference level social bias condition significantly differs from 0.5 bits (the input). We ran this model three times, with each social bias condition as the reference level. These analyses revealed that learners in the bias-for-case and the no-bias conditions produced linguistic systems that had significantly lower message uncertainty compared with the input (0.28 bits in the bias-for-case condition, β^ à 0:53, *t* à 2:49, *p* à 0:016; 0.14 bits in the no-bias matched condition, the β^à input 082).:93, message *t* à 4:37uncertainty , *p* < 0:0001)in . Learners their own in the productions bias-for-no-case (0:46condition, bits; β^àhowever, 0:04,

t àThese 0:22findings , *p* à 0: suggest that learners in the bias-for-no-case condition did not employ additional strategies to mitigate the increased uncertainty about semantic meaning compared with learners in the other social bias conditions.

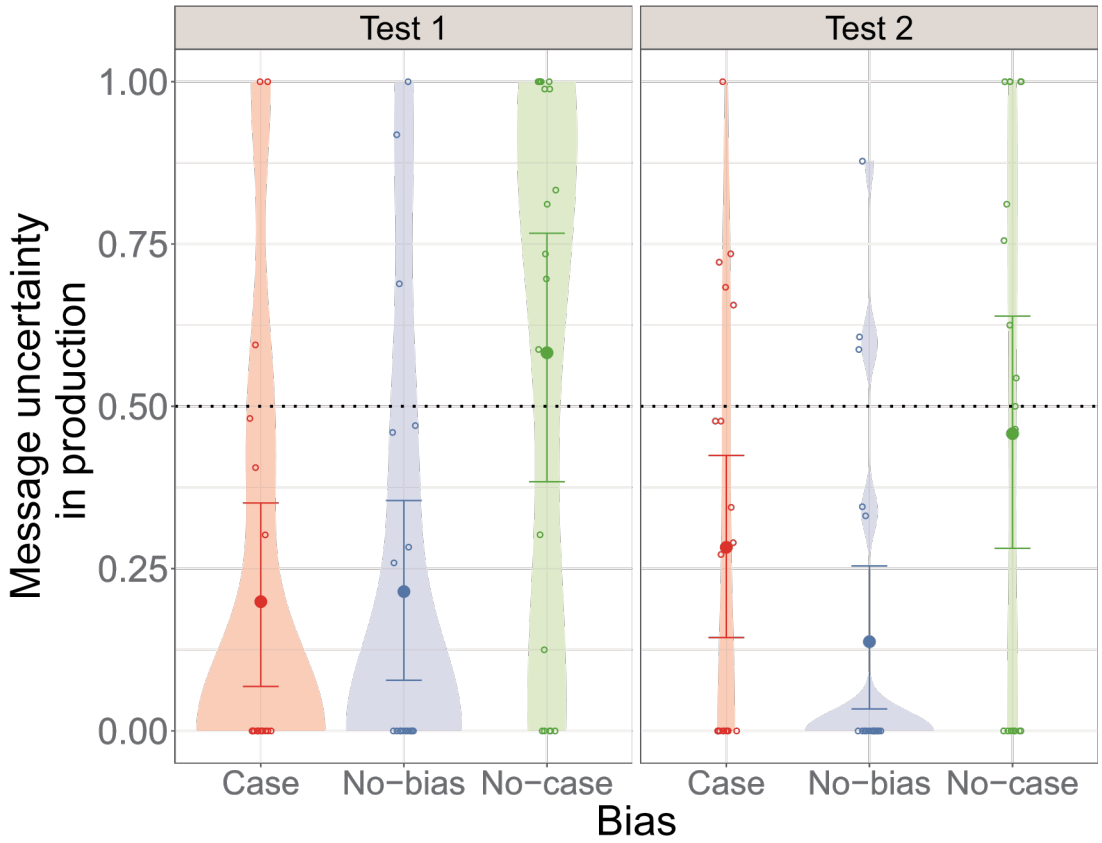


Figure 5. Uncertainty about the intended meaning in production by bias condition in Exp. 1. The dashed line represents the input proportion (same across social bias conditions). The large dots represent condition means. The small dots represent individual participant means. The error bars represent bootstrapped 95% confidence intervals.

Discussion of Experiment 1

We exposed participants to a language with two dialects that differed in whether they used case marking or not and manipulated whether participants were socially biased toward speakers of the dialect with case, the dialect with no case, or neither dialect in particular. We then compared learners' use of case marking and assessed the amount of message uncertainty in the linguistic systems they produced. The picture emerging from these results is that social biases and biases for efficient communication interact in shaping language change, via learners' grammatical choices, but do so in nonstraightforward ways. First, the social bias clearly influenced learners' use of case marking: Learners in the bias-for-no-case condition produced significantly less case marking compared with the other groups (30% case use in the bias-for-no-case condition, 63% and 65% case use in the bias-for-case and no-bias conditions, respectively, in the final production test in Experiment 1). Interestingly, a social bias *toward* speakers of the case dialect did not increase case use beyond the level in the no-bias condition, suggesting that the social bias did not override the preference to save production effort. The loss of case, however, came at a cost—learners in the bias-for-no-case condition produced systems that had higher uncertainty about the intended meaning compared with learners in the other conditions.

This is surprising because learners could have compensated for the increased uncertainty while still reducing case use. In particular, they could have fixed constituent order or conditioned case marking on it. Why did learners in the bias-for-no-case condition not employ such strategies? One possibility is that learners were simply insensitive to the increased uncertainty about the intended meaning created by case loss in the linguistic system. This possibility is rather unlikely, given the well-established findings from similar work suggesting that learners change miniature languages in ways that reduce such uncertainty (Fedzechkina et al., 2012, 2017). Another possibility is that the biases involved operate on different timelines. That is, the social cues might be sufficiently salient to exercise an early influence on the changes introduced by learners, but acquiring a sufficient grasp of the cues to semantic roles to compensate for increased uncertainty might take longer. This means that biases for efficient communication may not exercise their effect until later in the learning process. Indeed, there is some evidence that supports this idea. First, our learning task was hard for our participants, as indicated by high exclusion rates based on accuracy on the comprehension test. Second, work using similar paradigms typically shows that biases for efficient communication become most evident after substantial exposure to the miniature language (over several sessions; Fedzechkina et al., 2012, 2017). It is thus possible that the learning phase in our experiment was simply too short to give learners a chance to become comfortable enough with the novel language for the biases for efficient communication to exercise an effect on case use. We addressed this possibility in Experiment 2, in which we increased the amount of exposure.

Experiment 2

In Experiment 1, the exposure learners received was fairly short, given the complexity of the task. While we observed the influences of social biases on learners' grammatical choices, biases for efficient communication seemed to have less influence on learners' preferences, as evidenced by the fact that learners did not make up for the increased uncertainty caused by using less case in the bias-for-no-case condition compared to the other conditions. It is possible, however, that biases for efficient communication require a better command of the novel language (i.e., learners need to recognize that case and constituent order carry information about semantic meaning) and thus might require more exposure to the novel language to develop. In Experiment 2, we increased learners' exposure to the language from one 50-minute session (Experiment 1) to three 50-minute

sessions. Given evidence that sleep can enhance memory consolidation (Batterink & Paller, 2017), we also spread the learning sessions over three consecutive days.

Participants

In Experiment 2, we were interested specifically in what strategies (if any) learners in the bias-for-no-case condition would employ to reduce message uncertainty. The answer to this question required more complex statistical models than those used in Experiment 1. To ensure that we had adequate statistical power to detect the compensatory strategies of interest, we performed power simulations on the data from Experiment 1. This analysis revealed that while adequate power (80% for all effects) to detect the differences in case use between social bias conditions could be achieved with 20 participants per condition (as in Experiment 1), 40 successful learners per condition were needed to achieve adequate power to detect the specific strategies in case use to reduce message uncertainty. Thus, we set 40 successful learners as our recruitment target. After beginning recruitment, we observed that many participants did not return for all sessions of the experiment (see discussion of dropout rate below). We began recruiting more participants than our target in order to speed up data collection. In the end, more participants completed all three sessions than we anticipated. As such, the number of participants included in the analysis was uneven across conditions and slightly exceeded our target of 40 successful learners in each condition.

One hundred eighty-five participants completed all three sessions of the experiment via Prolific and FindingFive. All participants were self-reported monolingual native speakers of English with no known language disorders and at least 95% past approval on Prolific. Reflecting the difficulty of administering multiday experiments online, 112 participants dropped out after completing one or two sessions. Participants were paid $16.50 for completing all three sessions of the experiment (for a prorated hourly rate of $6.50). Participants who failed to complete all three sessions were paid for the sessions they had completed ($6.50 and $5 for the first and subsequent sessions, respectively; the reduced amount of payment over the sessions reflects the fact that participants take less time to complete the task as they become more proficient in the language).

Recruitment and execution of this study were approved by the Institutional Review Board of the University of Pennsylvania. Participant exclusion criteria were the same as in Experiment 1. Of the participants who completed all three days, 32 were excluded from the analysis for failing to adequately learn the language on the final comprehension test on the final day of the experiment, as defined in the Scoring and Exclusions section (13 participants in the bias-for-case, nine participants in the bias-for- no-case, ten participants in the no-bias condition). This left data from 153 participants (57 participants in the bias-for-case, 50 participants in the bias-for-no-case, 46 participants in the no-bias condition) for analysis.

Procedure

Each participant learned the same miniature language as in Experiment 1 and was assigned to one of the three social bias conditions used in Experiment 1. The experiment was administered in three sessions over three consecutive days with at least 24 hours between each pair of sessions. The procedure of Experiment 2 was identical to that of Experiment 1 (see Procedure section for Experiment 1 and Figure 3) with one exception: On the final (third) day, the two sentence production test blocks were administered back-to-back at the end of the experiment (instead of being separated by a sentence exposure and a comprehension test). This change in the procedure allowed us to double the amount of production data collected after participants had received all of

the novel language exposure, thus enabling us to more accurately estimate the strategies participants were using after successfully mastering the new language.

Results

In Experiment 2, we were primarily interested in learners' preferences in using the novel language after long exposure to it. Therefore, we assessed learners' performance in Experiment 2 based on their production data from Day 3 pooled across the two production tests. (As discussed above, the production tests in Experiment 2 were—unlike in Experiment 1—administered back-to-back at the end of Day 3 without additional exposure to the novel language between the two tests.) Before we turn to the discussion of case use and message uncertainty reduction in learners' production, we briefly discuss the accuracy of acquisition in this experiment.

Accuracy of acquisition

In sentence production on the final day of the experiment, participants made low levels of lexical errors (1.9% in the bias-for-case condition, 3.7% in the bias-for-no-case condition, 2.2% in the no-bias condition) and grammatical errors (2.9% in the bias-for-case condition, less than 1% in the bias-for-no -case and no-bias conditions; see Table 3). On the final comprehension test of Experiment 2, learners who were included in the analysis had an accuracy of over 98% on unambiguous (case marked) trials in all social bias conditions (see Table 4), suggesting that after three days of training, the overwhelming majority of participants had mastered the novel language well.

Case use in production

We first asked whether, after extensive training, the effect of social bias on case use persisted in learners' productions. That is, whether after three days of training, learners in the bias-for-no-case condition still used less case marking as a result of a social bias. To answer this question, we compared learners' case use across social bias conditions in sentence production pooled across both production tests on Day 3 in Experiment 2.

Table 3. Grammatical errors in production in Exp. 2

			Case errors		Constituent order errors	
Day	Test	Bias condition	Mean	95% CI	Mean	95% CI
Day 1	Test 1	Case	1.97%	0.76-3.51%	3.29%	0.11-7.02%
		No-bias	1.63%	0.27-3.80%	3.53%	0-8.28%
		No-case	1.63%	0.63-3.00%	1.38%	0-3.75%
	Test 2	Case	4.38%	1.32-8.01%	1.43%	0-3.94%
		No-bias	2.58%	0.81-4.76%	2.45%	0-7.07%
		No-case	1.25%	0.36-2.25%	0%	0-0%
Day 2	Test 1	Case	2.41%	0.88-4.28%	1.09%	0-3.07%
		No-bias	1.90%	0.54-3.80%	0%	0-0%

		No-case	0.38%	0-0.75%	0%	0-0%
	Test 2	Case	3.18%	0.77-6.25%	0.77%	0-1.97%
		No-bias	1.36%	0.13-2.85%	0%	0-0%
		No-case	0.38%	0-0.75%	0%	0-0%
Day 3	Test 1 & 2 (Pooled)	Case	1.97%	1.04-3.02%	0.88%	0-2.25%
		No-bias	0.61%	0.14-1.22%	0%	0-0%
		No-case	0.25%	0-0.56%	0%	0-0%

Table 4. Comprehension accuracy on final comprehension block each day in Exp. 2

		Accuracy	
Day	Bias condition	Mean	95% CI
Day 1	Case	91.45%	87.28-95.39%
	No-bias	90.22%	84.23-95.11%
	No-case	92.00%	88.00-95.50%
Day 2	Case	92.11%	86.40-96.93%
	No-bias	94.84%	91.03-97.83%
	No-case	96.25%	93.50-98.50%
Day 3	Case	98.25%	96.71-99.56%
	No-bias	98.10%	96.47-99.46%
	No-case	98.25%	96.75-99.50%

Specifically, we used mixed-effects logistic regression to predict case use from social bias condition (coded in the same way as in Experiment 1). The model contained the fullest converging random effects structure (random intercepts for participant and item, defined as object-noun).

This analysis revealed that learners' preferences in case use strongly depended on the social bias condition. Specifically, learners in the no-bias condition produced significantly less case compared to the learners in the bias-for-case condition (*β*^à case 2*:*18compared , *z* à 4*:*54to , learners *p* < 0*:*0001)in the , and no-bias learners condition in the

bias-for-no-case condition used significantly less

(*β*^àComparisons 2*:*9, *z* à of 5*:*learners' 33, *p* < 0case *:*0001use ; Figure to the 6input .) using mixed effects models with the same fixed and random effects structure and coding as in Experiment 1 further revealed that learners' case use preferences followed the input proportion of the alien dialect they were socially biased toward. Specifically, learners in the bias-for-case condition produced significantly more case compared with

the *p*condition <input 0:0001language produced); and as learners significantly a whole in (80%; the less *β*^ànо-bias case 7:54compared , *z* àcondition 7:51, *p*to <the matched 0:0001)input ; learners (20the %; input *β*^in àthe (527bias-for-no-case :73%; , *z β*^àà 07::4398, ,

z à 0:85, *p* à 0:397).

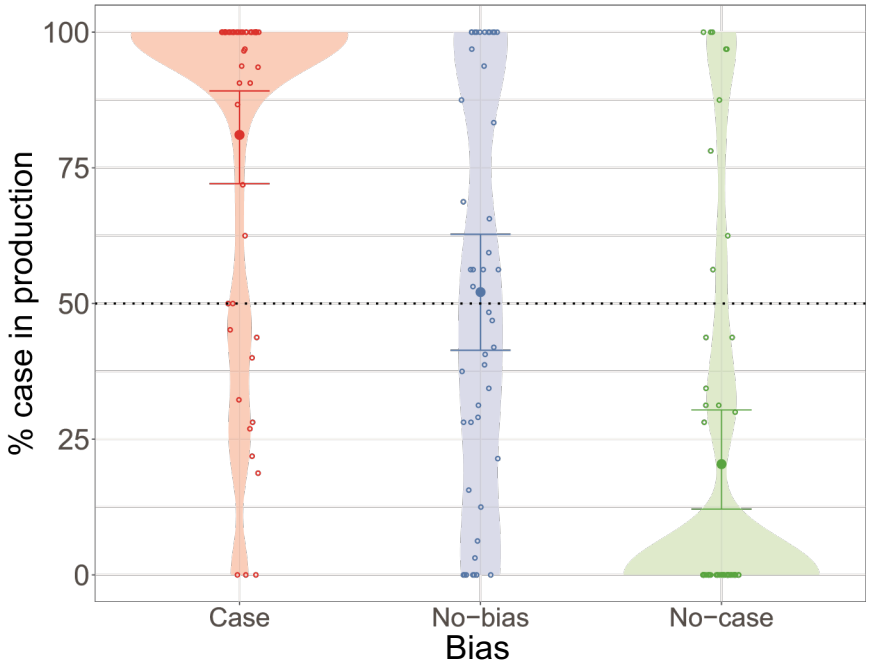


Figure 6. Case use in production by bias condition in Exp. 2. The dashed line represents the input proportion (same across social bias conditions). The large dots represent condition means. The small dots represent individual participants' means. The error bars represent bootstrapped 95% confidence intervals.

These findings suggest that the effects of the social bias were not fleeting and persisted even after prolonged exposure to the miniature language. We next asked whether, as a result of increased exposure to and better familiarity with the novel language, learners concomitantly made changes to the input language that would compensate for the increased uncertainty about the intended meaning brought about by dropping case.

Message uncertainty in production

As in Experiment 1, we calculated the conditional entropy in production on the final day of training (Figure 7). We then conducted a linear regression analysis to predict conditional entropy in production from social bias condition (coded the same way as in case use in production). Again, we found that our residuals were not normally distributed. For Experiment 2, bestNormalize determined that the best transformation was ordered quantile normalization (Peterson & Cavanaugh, 2020). The analysis with the transformed data revealed that learners in the no-bias condition of Experiment 2 produced linguistic systems with significantly more uncertainty about the intended meaning com-

pared to learners in the bias-for-case condition (*β*^à 0:that 20, *t*did à 3not *:*27, significantly *p* à 0:001), diand ffer learners in amount in the of

bias-for-no-case condition produced linguistic systems uncertainty from the no-bias baseline (*β*^à

0:09the , *t* àsame 1:39, linear *p* à 0:17).regression model structure as in

We performed input comparisons using
Experiment 1 on the transformed conditional entropy data from Experiment 2. These comparisons revealed that while learners in the bias-for-case and no-bias conditions produced linguistic systems that had significantly lower message uncertainty compared to the input (0.086 bits in the bias-for-case condition; *β*^à 0:4009)*:*14, , learners *t* à 9:39in , the *p* <bias-for-no-case 0:0001; and 0.27condition bits in produced the no-bias linguistic condition; systems *β*^àthat 0:did 36,

t à 2:63, *p* à

not significantly differ from the input (0.5 bits) message uncertainty (0.4 bits in the bias-for-no-case condition; *β*^ à 0:09, *t* à 0:73, *p* à 0:47).

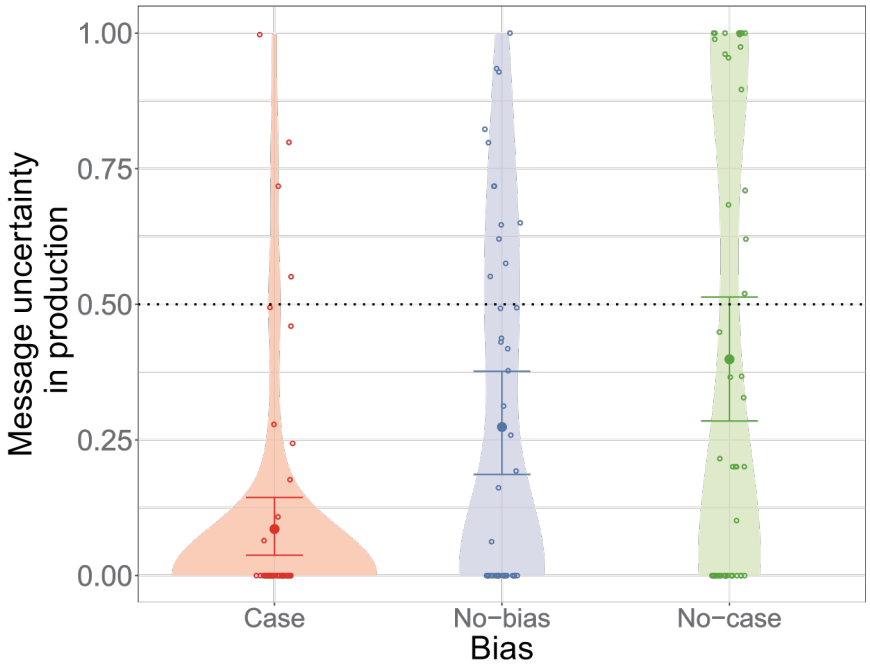


Figure [6]. Message uncertainty in production by bias condition in Exp. 2. The dashed line represents the input proportion (same across social bias conditions). The large dots represent condition means. The small dots represent individual participant means. The error bars represent bootstrapped 95% confidence intervals.

In other words, prolonged exposure to the input did not lead to a significant reduction in the message uncertainty caused by the loss of case. This suggests that if learners in the bias-for-no-case condition were attempting to implement any strategies to mitigate the effects of dropping case,

[6] We include graphs of the equivalent data from Experiment 1 (Figures B1 and B2 in Appendix B), but do not provide statistical analysis because there was not adequate power for models investigating specific strategies, as explained in the Participants section.

they were subtle at best. But was there any evidence that participants were attempting to implement such mitigation strategies (if only subtly)? Given the grammar of the miniature language, participants could have chosen to use one constituent order with higher frequency than the other, or they could have conditioned case use on constituent order and used it more frequently with one constituent order variant. Both strategies would reduce message uncertainty. To investigate where either strategy was employed, we performed additional mixed effects analyses on the production data in Experiment 2. These analyses revealed that learners produced SOV order equally often across all bias conditions (no- bias vs. bias-for-case condition: β^ à 0:03, z à 0:21, p à 0:83; bias-for-no-case vs. no-bias condition:

case in OSV
ing more β^àprior 0:05findings , z à 0on order :28case , $p$compared àuse 0:77in ; this Figure with paradigm SOV 8). Additionally, order (Fedzechkina overall (we βàet found 0al., :302017, that z à). 3learners There :63, p àwere used 0:0002)no significantly significant , replicat-

interactions between sentence constituent order and any of the social bias conditions (no-bias vs. bias-

for-case condition condition condition case ⇥ sentence on OSV ⇥ sentence order: constituent β^

order: 0:079β^à, z0:à04not , z1di:à21ffer 0, :p791across à,

0p:22)àsocial , 0:suggesting 429bias ; bias-for-no-case

conditions.that the [7] preference These vs. no-bias results to à order

did

indicate that, while there was variation in case use across conditions—which was driven by social biases—there was no evidence of variation in learners' use of other grammatical devices, even when it might have mitigated the increase in message uncertainty resulting from low use of case marking.

Discussion of Experiment 2

We replicated Experiment 1 with a longer, three-day, exposure. We found that, after longer exposure to the novel language, case use in production came to reflect the input proportion of the dialect to which learners had been socially biased, suggesting that the social biases had a persistent, nonfleeting effect on case use. Dropping case still came at a cost, however. While the message uncertainty in the bias-for-no-case condition was not significantly higher than in the no-bias condition, it was also no lower than the input level of 0.5; this was the only condition in which this

was so. In other words, while the output language in this condition had adapted to the social biases, there was still a high chance of miscommunication with regard to semantic meaning, in spite of the prolonged training, and no evidence that participants responded by changing the language in ways that would reduce the potential for miscommunication.

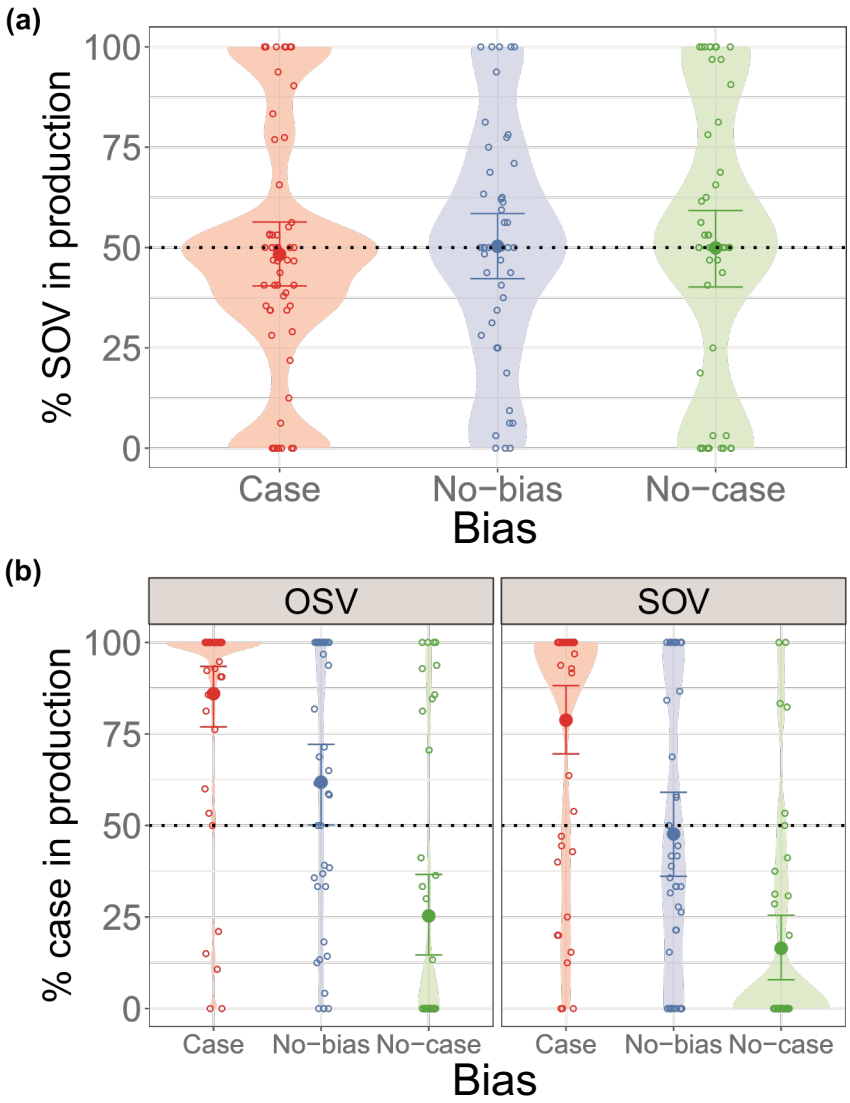


Figure 8. Constituent order (top panel) and conditioning of case on constituent order (bottom panel) in production by bias condition in Exp. 2. The dashed lines represent the input proportion (same across social bias conditions). The large dots represent condition means. The small dots represent individual participant means. The error bars represent bootstrapped 95% confidence intervals.

General discussion

Across two experiments, we exposed participants to a miniature language with flexible constituent order and two dialects, one that employed case marking in all sentences and one that contained no case marking at all. In the first experiment, which took place over a single session, we varied the instructions in a between-subjects design to bias participants (a) toward the dialect with no case, (b) toward the dialect with case, or (c) toward neither dialect in particular. The second experiment replicated the findings from the first experiment with extended exposure to the new language over three days. In both experiments we measured case use in participants' own productions of the novel language as a result of the social bias and assessed the message uncertainty of the linguistic systems

produced by participants. In both experiments we found that social biases played an important role in language change, regardless of the consequences for robust communication of semantic meaning.

In particular, we observed clear influences of social biases on participants' case use in production, with learners in the bias-for-no-case condition producing significantly less case marking compared with all other social bias conditions. This preference was evident after short (single-session) exposure to the miniature language and did not change with longer (three-session) exposure. These findings conceptually replicate and extend prior work by Roberts and Fedzechkina (2018), who found that learners exposed to a language with fixed (i.e., informative) constituent order and redundant case marking (which required effort to produce) maintained case marking longer if there was a social bias to use it. However, the current study differs from Roberts and Fedzechkina's work in a very important respect—the informativity of case marking. In Roberts and Fedzechkina's study, case marking was redundant, so the uncertainty about the intended meaning remained extremely low, regardless of the presence or absence of case marking. In the current experiments, constituent order was uninformative and case marking carried important information about the intended meaning. The loss of case marking as a result of a social bias in our experiments, therefore, came at a cost of increasing message uncertainty. In other words, our study goes beyond the earlier work in showing that learners follow a social bias even if it leads to linguistic systems that are less desirable with regard to the efficient communication of semantic meaning, being *less robust* in distinguishing subject from object.

However, communicating social meaning efficiently need not be at odds with efficiently communicating semantic meaning. This is true even if communicating social meaning involves dropping semantically meaningful case. Participants could have mitigated the message uncertainty arising from the reduction in case use by making other changes to the grammar, such as fixing word order or conditioning case use on word order. We found little evidence of message uncertainty reduction in either experiment, however. In Experiment 1, learners of the bias-for-no-case condition produced linguistic systems that had significantly higher uncertainty compared to the no-bias baseline. Experiment 2 involved prolonged training, which earlier work had suggested might give participants greater opportunity to make such changes (Fedzechkina et al., 2012, 2017). We did not, however, find convincing evidence for message uncertainty reduction even after longer exposure. First, overall message uncertainty in the bias-for-no-case condition was not significantly below the input level of 0.5 bits. Second, constituent order, one pathway for uncertainty reduction in our experiment, did not differ across social bias conditions. Third, while we found evidence in Experiment 2 that participants had conditioned case use on constituent order to some extent (a means of mitigating message uncertainty by making the absence of case informative), this did not differ between conditions. There was thus no evidence that participants were particularly motivated by the loss of case in the bias- for-no-case condition to mitigate the message uncertainty that resulted from it.

This raises the question of why such mitigation did not occur in our experiment. One possibility is that the biases involved in our study might interact in rather complex ways that only play out fully over more than one generation. Thus, an interesting focus for future work would be an iterated learning study where the output of one generation of learners is passed as the input to the next generation, a process that is known to amplify weak biases (cf. Kirby et al., 2014). Another possibility is that some grammatical devices are more susceptible to distributional changes introduced by the learners than others. The case marker was represented in our experiment by a one-syllable suffix -*dak* that followed the noun it modified and was likely less salient in the input than the content nouns. It is thus possible that the changes to constituent order distribution were associated with

larger perceptual changes for our participants and therefore were avoided. This possibility is consistent with prior work using similar miniature languages that found no deviations from the constituent order distributions while finding deviations from case marking distributions (Fedzechkina & Jaeger, 2020; Fedzechkina et al., 2017).

Yet another possibility concerns the inclusion and manipulation of specific addressees during sentence production. For simplicity's sake, the current study did not involve an explicitly defined interlocutor. Nor (in order to avoid introducing any extra bias) did production trials involve feedback that penalized sentences with high message uncertainty. This does not mean that no communicative pressures were present in our setup. A large body of evidence suggests that language users tailor their utterances not only for actual interlocutors but also for *potential* conversational partners (such as noninteracting addressees, overhearers, or even imagined or expected addressees; Ferreira & Dell, 2000, Clark & Schaefer, 1992, Wade & Roberts, 2020). Furthermore, it has been widely shown that the biases we termed "biases for efficient communication" (i.e., biases to reduce message uncertainty and production effort) are not restricted to situations in which people interact with one another; they also operate in contexts in which individuals are encoding messages for themselves, with no interlocutors present, both in natural language use (Kurumada & Jaeger, 2015; Levi, Bicknell, Slattery, & Rayner, 2009; Mahowald, Fedorenko, Piantadosi, & Gibson, 2013) and in miniature language paradigms similar to ours (Fedzechkina et al., 2012; Kurumada & Grimm, 2019). At the same time, however, long-standing work on experimental communication games strongly suggests that feedback and communicative interaction with an actual partner do play a role in speakers' utterance design (Fay, Walker, Swoboda, & Garrod, 2018; Kanwal, Smith, Culbertson, & Kirby, 2017; Schober & Clark, 1989), and our task lacked many of the pragmatic cues that are present in a natural communicative interaction. It therefore seems likely that the presence of a conversational partner (whether human or simulated, as in Buz, Tanenhaus, & Jaeger, 2016) would boost the effect of biases for efficient communication, leading to a greater reduction in message uncertainty. Another obvious focus for future work is manipulating the identity of perceived interlocutors (e.g., as belonging to one or another group of aliens; cf. Sneller & Roberts, 2018; Wade & Roberts, 2020) or their linguistic behavior (e.g., whether they use linguistic markers variably or categorically; cf. Fehér, Ritt, & Smith, 2019). We would expect this to lead to more complex, or perhaps differently structured, interactions between the social and nonsocial communicative pressures, potentially reducing or boosting their effects differentially, depending on the interlocutor involved.

Another question that remains unanswered in the current work concerns the mechanism through which the social bias operates. One possibility is that by simply mentioning a particular dialect in the instructions, we drew participants' attention to it, which made them learn the case distribution in this dialect more accurately than in the other dialect. It is hard to identify from our data whether this was the case, however. Since our experiment was intended to test participants' *use* of grammatical devices (and successful learning of the function of these devices is a prerequisite for use), our paradigm was designed to achieve highly successful learning. Indeed, learners in all conditions achieved high comprehension and production accuracy, which does not allow us to distinguish learning patterns in the different social bias conditions. Nonetheless, prior work in this paradigm suggests that it is unlikely our results arose due to differential attention during learning. Specifically, Roberts and Fedzechkina (2018) included a condition that biased participants *against* the case dialect. If a simple mention of the case dialect increased learners' attention to it, we would expect learners in this condition to use more case. However, this did not occur—instead, learners in the bias-against-case condition dropped case to the same degree as learners biased in favor of

the no-case dialect, suggesting that the instructions induced a social bias toward the different alien species rather than merely directing attention to their dialect.

Another possibility is that our experimental setup provided different incentives for communicating social versus semantic meaning. While our instructions specifically encouraged learners to feel positively inclined toward a particular group or groups of aliens, they did not specifically instruct participants to convey the intended message in such a way that an alien speaker could understand it. A way to explore this possibility in future work would be to manipulate the relative importance of communicating social versus semantic meaning in the same task. For instance, success in the game could be stated to depend on impressing the aliens or reliably communicating information to them (or some combination of both). If our findings are driven by the perceived importance of different kinds of information, we would expect learners to prioritize different types of information, depending on how successful communication is defined. Along similar lines, it would be interesting to include interlocutors of different species in the production phase; this would allow us to investigate the extent to which participants respond strategically to the same bias, depending on context (cf. Sneller & Roberts, 2018; Wade & Roberts, 2020).

A related limitation concerns the fact that our input languages involved categorical differences between the dialects (one of which employed no case at all, while the other employed it 100% of the time). This is not typical of natural language variation, which has long been known to be characterized by more gradient patterns, with individual speakers using more than one variant, conditioned on such factors as register (Roberts & Sneller, 2020; Weinreich, Labov, & Herzog, 1968). It is unlikely, for instance, that many English speakers who use *whom* or *yous* use it all the time across all registers. Such graded variation could be incorporated into input languages in the experimental paradigm we employed (cf. Lai, Rácz, & Roberts, 2020) without reducing the overall level of case in the input language. Doing so not only would increase the ecological validity of our experiment but also might result in different patterns of change being introduced by learners into the input language. Specifically, more graded variation could potentially reduce case loss in the output language: A social bias favoring a dialect with relatively lower levels of case establishes a less categorical target than a social bias favoring a dialect with no case whatsoever. This could give participants an opportunity to be socially consistent while using case more flexibly, which could make them more likely to condition it on constituent order, thus creating linguistic systems with low message uncertainty.

Such limitations aside, our work fits in well with existing work on related questions. In the absence of a social bias (i.e., in the no-bias condition), learners retained informative case marking, producing it above (Experiment 1) or at the input proportion (Experiment 2). These results are consistent with a growing body of information-theoretic work on the emergence of cross-linguistic trade-offs in cues to semantic roles. In particular, participants in our experiment who were presented with a language that had flexible (i.e., uninformative) constituent order maintained an additional cue to semantic meaning (case) to reduce uncertainty about the intended meaning, and they did so in spite of the additional effort cost of producing case. This is consistent both with cross-linguistic patterns in natural language (Koplenig et al., 2017) and with prior work using a similar experimental paradigm (e.g., Fedzechkina & Jaeger, 2020; Fedzechkina et al., 2017; Hall Hartley & Fedzechkina, 2020).

In conclusion, we extended an established experimental paradigm to study the interaction of social biases and biases for efficient communication, shining new light on the complex ways in which they jointly shape language change. We found not only that social biases modulate the role of biases for efficient communication but also that they can lead to linguistic systems that are less

communicatively efficient, such as systems with *increased*, rather than *reduced*, message uncertainty. We also laid the groundwork for a variety of future extensions of the paradigm.

Acknowledgments

We thank Aja Altenhof for help with participant recruitment, as well as Charlie Torres and Vanessa Nieto for help with stimuli creation. The second experiment was supported by a grant from the University of Pennsylvania University Research Fund. The third author was also supported by the National Science Foundation (grant number 1946882).

Data availability

The data that support the findings of this study are openly available in an Open Science Foundation repository at https://osf. io/hb9dc/

Disclosure Statement

No potential conflict of interest was reported by the authors.

ORCID

Gareth Roberts http://orcid.org/0000-0002-6662-6829

References

Ackerman, F., & Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, *89* (3), 429–464.

Batterink, L. J., & Paller, K. A. (2017). Sleep-based memory processing facilitates grammatical generalization: Evidence from targeted memory reactivation. *Brain and Language*, *167*, 83–93.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 1–20.

Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory & Language*, *89*, 68–86.

Clark, H. H., & Schaefer, E. F. (1992). Dealing with overhearers. In H. H. Clark (Ed.), *Arenas of language use*. Chicago: University of Chicago Press.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*, 306–329. doi: 10.1016/j.cognition.2011.10.017

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12*(4), 453–476.

Fay, N., Walker, B., Swoboda, N., & Garrod, S. (2018). How to create shared symbols. *Cognitive Science*, *42*, 241–269.

Fedzechkina, M., Chu, B., & Jaeger, T. (2018). Human information processing shapes language change. *Psychological Science*, *29*(1), 72–82.

Fedzechkina, M., & Jaeger, T. F. (2020). Production efficiency can cause grammatical change: Learners deviate from the input to better balance efficiency against robust message transmission. *Cognition*, *196*, 104115.

Fedzechkina, M., Jaeger, T., & Newport, E. (2012). Language learners restructure their input to facilitate efficient communication. *Proc Natl Acad Sci USA*, *109*(44), 17897–17902. doi: 10.1073/pnas.1215776109

Fedzechkina, M., Newport, E., & Jaeger, T. (2016). The miniature artificial language learning paradigm as a complement to typological data [Book Section]. In L. Ortega, A. Tyler, H. Park, & M. Uno (Eds.), *The usage-based study of language learning and multilingualism*. Georgetown: GUP.

Fedzechkina, M., Newport, E., & Jaeger, T. (2017). Balancing effort and information transmission during language acquisition: Evidence from word order and case-marking. *Cognitive Science*, *41*(2), 416–446 . doi: 10.1111/cogs.12346

Fehér, O., Ritt, N., & Smith, K. (2019). Asymmetric accommodation during interaction leads to the regularisation of linguistic variants. *Journal of Memory and Language*, *109*, 104036.

Ferreira, V., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology 40*(4), 296–340.

Finding Five Corporation. (2019). *FindingFive: A web platform for creating, running, and managing your studies in one place*. Retrieved from https://www.findingfive.com

Frank, A., & Jaeger, T. (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. In *Proceedings of the 30th annual meeting of the Cognitive Science Society,* 933–938.

Hall Hartley, L., & Fedzechkina, M. (2020). Learners' bias to balance production effort against message uncertainty is independent of their native language. In *Proceedings of the 42nd annual conference of the Cognitive Science Society*.

Hasegawa, A. (2011). *The semantics and pragmatics of Japanese focus particles* (Doctoral dissertation, State University of New York at Buffalo). https://arts-sciences.buffalo.edu/content/dam/arts-sciences/linguistics/AlumniDissertations/Hasegawa%20dissertation.pdf

Hudson Kam, C., & Newport, E. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*(1), 30–66.

Jäger, G. (2007). Evolutionary game theory and typology: A case study. *Language*, 74–109.

Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. (2017). Zipf's law of abbreviation and the principle of least effort: Language users optimise a miniature lexicon for efficient communication. *Cognition*, *165*, 45–52.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114.

Koplenig, A., Meyer, P., Wolfer, S., & Mueller-Spitzer, C. (2017). The statistical trade-off between word order and word structure–large-scale evidence for the principle of least effort. *PloS one*, *12*(3).

Kurumada, C., & Grimm, S. (2019). Predictability of meaning in grammatical encoding: Optional plural marking. *Cognition* 191. 103953. https://doi.org/10.1016/j.cognition.2019.04.022

Kurumada, C., & Jaeger, T. (2015). Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language*, *83*, 152–178. doi: 10.1016/j.jml.2015.03.003

Labov, W. (2001). *Principles of linguistic change volume 2: Social factors* (Vol. 29). Hoboken, NJ: Blackwell.

Lai, W., Rácz, P., & Roberts, G. (2020). Experience with a linguistic variant affects the acquisition of its sociolinguistic meaning: An alien-language-learning experiment. *Cognitive Science*, *44*(4), e12832.

Lasnik, H., & Sobin, N. (2000). The who/whom puzzle: On the preservation of an archaic feature. *Natural Language & Linguistic Theory*, *18*(2), 343–371.

Launey, M. (2011). *An introduction to classical Nahuatl*. Cambridge: Cambridge University Press.

Levi, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences, USA*(106), 21086–21090.

Levshina, N. (2021). Cross-linguistic trade-offs and causal relationships between cues to grammatical subject and object, and the problem of efficiency-related explanations. *Frontiers in Psychology*(12), 2791.

Mahowald, K., Fedorenko, E., Piantadosi, S., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*(126), 313–218.

Peterson, R. A. (2021). Finding Optimal Normalizing Transformations via bestNormalize. *The R Journal*, *13*(1), 310–329. doi: 10.32614/RJ-2021-041

Peterson, R. A., & Cavanaugh, J. E. (2020). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*, *47*(13–15), 2312–2327. doi: 10.1080/02664763.2019.1630372

Preston, D. R. (1998). They speak really bad English down South and in New York City. In L. Bauer & P. Trudgill (Eds.), *Language myths*. London: Penguin.

Preston, D. R. (1999). A language attitude approach to the perception of regional variety. In *Handbook of perceptual dialectology*. Amsterdam: John Benjamins.

Preston, D. R. (2015). The silliness of the standard. *Representaciones. Revista de Estudios sobre Representaciones en Arte, Ciencia y Filosofía*, *11*(2).

Puskás, G. (2000). *Word order in Hungarian: The syntax of Ā-positions*. Amsterdam: John Benjamins.

Roberts, G., & Fedzechkina, M. (2018). Social biases modulate the loss of redundant forms in the cultural evolution of language. *Cognition*, *171*, 194–201.

Roberts, G., & Sneller, B. (2020). Empirical foundations for an integrated study of language evolution. *Language Dynamics and Change*, *10*, 188–229.

Sapir, E. (1921). *Language: An introduction to the study of speech*. New York: Harcourt, Brace.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*(2), 211–232.

Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, *116*(3), 444–449. doi: 10.1016/j.cognition.2010.06.004

Sneller, B., & Roberts, G. (2018). Why some behaviors spread while others don't: A laboratory simulation of dialect contact. *Cognition*, *170*, 298–311.

Stevens, J. S., & Roberts, G. (2019). Noise, economy, and the emergence of information structure in a laboratory language. *Cognitive Science*, *43*(2), e12717.

Tallerman, M. (2006). The syntax of Welsh "direct object mutation" revisited. *Lingua*, *116*(11), 1750–1776.

Van Everbroeck, E. (2003). Language type frequency and learnability from a connectionist perspective. *Linguistic Typology*, *7*(1), 1–50. doi: 10.1515/lity.2003.011

Wade, L., & Roberts, G. (2020). Linguistic convergence to observed versus expected behavior in an alien- language map task. *Cognitive Science*, *44*(4), e12829.

Weber, A., Grice, M., & Crocker, M. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, *99*(2), B63–B72.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics* (pp. 95–195). Austin: University of Texas Press.

Appendix A. Full model results

Table A1. Results from the generalized linear mixed effects model predicting the presence of case marking from social bias condition, production test block, and their interaction in Exp. 1.

Fixed effect	Estimate	SE	*z*-score	*p*-value
(Intercept)	-0.035	0.625	-0.055	0.956
No-bias vs. bias-for-case condition	-0.304	0.495	-0.614	0.539
Bias-for-no-case vs. no-bias condition	-1.659	0.531	-3.123	**0.002**
Test block 2	0.867	0.298	2.903	**0.003**
No-bias vs. bias-for-case condition: Test block 2	0.010	0.233	0.044	0.965
Bias-for-no-case vs. no-bias condition: Test block 2	-0.250	0.289	-0.867	0.386

Table A2. Results from the linear regression model predicting conditional entropy (exponentially transformed) from social bias condition, production test block, and their interaction in Exp. 1.

Fixed effect	Estimate	SE	*t*-value	*p*-value
(Intercept)	0.002	0.086	0.019	0.9850
No-bias vs. bias-for-case condition	-0.051	0.070	-0.733	0.4650
Bias-for-no-case vs. no-bias condition	0.304	0.070	4.339	< **0.0001**
Test block 2	-0.065	0.086	-0.756	0.4510
No-bias vs. bias-for-case condition: Test block 2	-0.060	0.070	-0.852	0.3960
Bias-for-no-case vs. no-bias condition: Test block 2	-0.026	0.070	-0.369	0.7130

Table A3. Results from the generalized linear mixed effects model predicting the presence of case marking from social bias condition in Exp. 2.

Fixed effect	Estimate	SE	*z*-score	*p*-value
(Intercept)	0.262	0.523	0.501	0.6160
No-bias vs. bias-for-case condition	-2.188	0.481	-4.546	< **0.0001**
Bias-for-no-case vs. no-bias condition	-2.901	0.544	-5.333	< **0.0001**

Table A4. Results from the linear regression model predicting conditional entropy (ordered quantile norm transformed) from social bias condition in Exp. 2.

Fixed effect	Estimate	SE	*t*-value	*p*-value
(Intercept)	0.025	0.076	0.336	0.740
No-bias vs. bias-for-case condition	0.202	0.062	3.274	**0.001**
Bias-for-no-case vs. no-bias condition	0.088	0.064	1.390	0.166

Table A5. Results from the generalized linear mixed effects model predicting word order use from social bias condition in Exp. 2.

Fixed effect	Estimate	SE	*z*-score	*p*-value
(Intercept)	-0.027	0.226	-0.119	0.905
No-bias vs. bias-for-case condition	0.038	0.181	0.207	0.836
Bias-for-no-case vs. no-bias condition	-0.054	0.188	-0.288	0.773

Table A6. Results from the generalized linear mixed effects model predicting the presence of case from social bias condition, word order used, and their interaction in Exp. 2.

Fixed effect	Estimate	SE	*z*-score	*p*-value
(Intercept)	0.258	0.520	0.495	0.6200
No-bias vs. bias-for-case condition	-2.172	0.480	-4.524	< **0.0001**
Bias-for-no-case vs. no-bias condition	-2.882	0.542	-5.314	< **0.0001**
OSV word order	0.302	0.083	3.638	**0.0002**
No-bias vs. bias-for-case condition: OSV word order	0.050	0.063	0.791	0.4290

Bias-for-no-case vs. no-bias condition: OSV word order	-0.080	0.066	-1.218	0.2230

Appendix B. Mitigation strategy analysis in Experiment 1

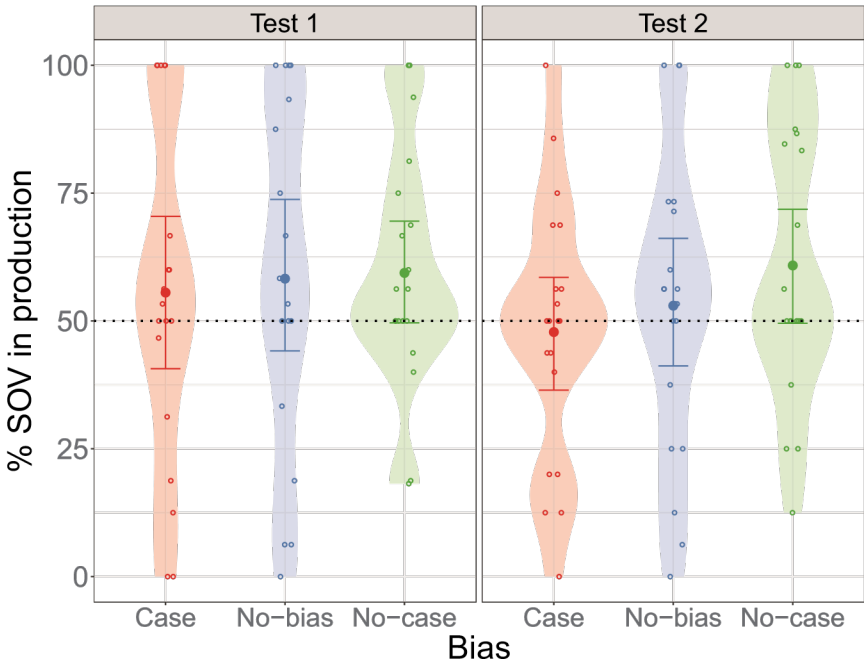


Figure B1. Constituent order in production by bias condition in Exp. 1. The dashed lines represent the input proportion (same across social bias conditions). The large dots represent condition means. The small dots represent individual participant means. The error bars represent bootstrapped 95% confidence intervals.

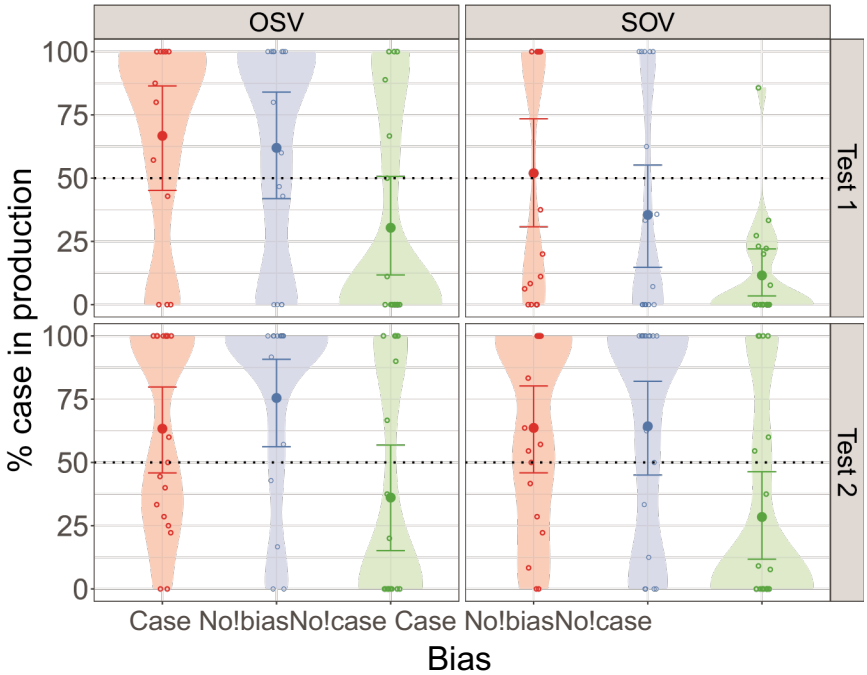


Figure B2. Conditioning of case use on constituent order in production by bias condition in Exp. 1. The dashed lines represent the input proportion (same across social bias conditions). The large dots represent condition means. The small dots represent individual participant means. The error bars represent bootstrapped 95% confidence intervals.