Cognition 196 (2020) 104115

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

# Production efficiency can cause grammatical change: Learners deviate from the input to better balance efficiency against robust message transmission

Masha Fedzechkina[a,b,c,*], T. Florian Jaeger[d,e]

[a] Department of Linguistics, University of Arizona, Tucson, AZ, 85721, USA
[b] Graduate Interdisciplinary Program in Cognitive Science, University of Arizona, Tucson, AZ, 85721, USA
[c] Graduate Interdisciplinary Program in Second Language Acquisition and Teaching, University of Arizona, Tucson, AZ, 85721, USA
[d] Department of Brain & Cognitive Sciences, University of Rochester, Rochester, NY, 14627-0268, USA
[e] Department of Computer Science, University of Rochester, Rochester, NY, 14627-0268, USA

ARTICLE INFO

*Keywords:*
Communicative efficiency
Miniature language learning
Information theory
Language universals
Language acquisition

ABSTRACT

The idea that human languages have properties suitable for efficient communication has permeated linguistic theorizing. Indirect correlational support for this idea has come from cross-linguistic synchronic and diachronic data. However, direct causal tests have been lacking. We directly test whether biases operating during language learning can *cause* learners to deviate from the input they receive towards output languages that better balance production efficiency against robust message transmission. We employ miniature language learning experiments to address this question for a well-documented cross-linguistic correlation between constituent order flexibility and the presence of case marking in a language. Participants were exposed to novel miniature languages that had *optional* case marking and either fixed or flexible constituent order. Between participants, we manipulated the amount of time and effort associated with the production of case marking. We find that learners introduced the cross-linguistically observed trade-off between case marking and constituent order flexibility into their output languages. Critically, learners only did so when case-marked nouns required additional effort compared to non-case-marked nouns. Thus, the present study suggests that even abstract grammatical properties of languages can be shaped by a balance between production efficiency and robust message transmission.

## 1. Introduction

In his seminal work, Zipf (1949) popularized the idea that languages have properties that make them particularly suited for human communication. He suggested in particular that human goal-directed behaviors including language are organized according to a single principle of least effort that seeks to minimize the expected total work expenditure required to achieve the goal. On this account, efficient communicative systems adopted by a community of speakers need to strike a balance between two competing pressures—the preferences of the speaker to conserve production efficiency (or the *force of unification* in Zipf's terminology) and the preferences of the listener to successfully recover the intended meaning (or the *force of diversification*).[1] The joint influence of these forces can explain basic properties of natural language lexicons, such as the relationship between word length and frequency: Words that are on average highly frequent in their context tend to be short as it saves production effort for the speaker without hindering word recognition for listeners as frequent words are more expected.

It is hard to overestimate the influence that this idea has had in language sciences over the years. The hypothesis that languages represent an efficient balance of two competing pressures has recurred in various functionalist accounts as competition between the forces of iconicity and economy (Aissen, 2003; Bates & MacWhinney, 1982; Croft, 2003; Givón, 1985; Hockett, 1958), clarity and ease (Slobin, 1977), or efficiency and complexity (Hawkins, 2004). This balance is invoked in explanations for a broad variety of phenomena in natural languages ranging from the structure of the lexicon to patterns in word order and morphology (Cohen Priva, 2015; Currie-Hall, Hume, Jaeger, & Wedel, 2018; Maurits, Perfors, & Navarro, 2010; Wedel, Kaplan, & Jackson, 2013).

The prediction that words that are on average more expected tend to have shorter phonological forms has now been validated across a variety of languages (Manin, 2006; Meylan & Griffiths, 2017;

* Corresponding author at: Department of Linguistics, University of Arizona, 1103 E University Blvd, Tucson, AZ, 85721, USA.
*E-mail address:* mfedzech@email.arizona.edu (M. Fedzechkina).

[1] The idea of competing pressures popularized by Zipf can be traced at least as far back as von der Gabelentz (1901) who spoke about a balance between the pressures for ease (*Bequemlichkeitsstreben*) and clarity (*Deutlichkeitsstreben*).

https://doi.org/10.1016/j.cognition.2019.104115
Received 18 February 2018; Received in revised form 21 October 2019; Accepted 23 October 2019
0010-0277/ 

Piantadosi, Tily, & Gibson, 2011). The same balance between expectedness and phonological form has also been found to shape the lexicon diachronically, so that the expectedness of a word at one historical time is predictive of the word's phonetic form at a later time (Soskuthy & Hay, 2017). However, alternative explanations for these correlational patterns have also been proposed. For example, a lexicon generated by random typing tends to have shorter frequent 'words', suggesting that the properties often attributed to constraints on efficient communication may arise due to an unrelated (and as of yet unknown) statistical process (Ferrer i Cancho & Moscoso del Prado Martin, 2013; for responses and related discussion see Piantadosi, 2014; Piantadosi, Tily, & Gibson, 2013). Since the bulk of support for Zipf's hypothesis comes from cross-linguistic correlational data, such alternative explanations substantially weaken the support for Zipf's hypothesis.

Recent work has begun to put Zipf's hypothesis to a stronger test. This work—like the present study—asks whether the hypothesized balance between production efficiency and robust message transmission can *cause* language change over time. One approach to this question is through miniature artificial language learning, in which participants learn and use novel miniature languages within a laboratory setting (Gomez & Gerken, 1999; Saffran, Aslin, & Newport, 1996). A growing number of studies within this paradigm has found parallels between patterns in cross-linguistic diversity and biases operating during the acquisition of miniature languages. In these studies, cross-linguistically frequent patterns are typically learned more easily or are preferred over less frequent alternatives (Culbertson, Smolensky, & Legendre, 2012; Fedzechkina, Chu, & Jaeger, 2017; Finley & Badecker, 2008; Hupp, Sloutsky, & Culicover, 2009; Kirby, Tamariz, Cornish, & Smith, 2015; Smith & Wonnacott, 2010; St Clair, Monaghan, & Ramscar, 2009). But it was not until recently, that some studies began to employ this paradigm to directly study the joint influence of production efficiency and robust message transmission on the acquisition of the lexicon (Kanwal, Smith, Culbertson, & Kirby, 2017) and grammar (Fedzechkina, Jaeger, & Newport, 2012; Fedzechkina, Newport, & Jaeger, 2016).

For example, Kanwal and colleagues used a miniature language learning paradigm to show that the relationship between word length and frequency in the lexicon emerges as a result of *joint* influence of production efficiency and robust message transmission (Kanwal et al., 2017). In their experiment, participants learned the names for two novel objects. Each of the two objects could be referred to by either a short or a long word form (e.g., *zopekil* and *zop*). Critically, the short form (*zop*) was shared between the two objects, making it ambiguous. Kanwal and colleagues found that learners produced the short form for the more frequent meaning, and the long form for the less frequent meaning, but did so only under the joint pressure of *both* production efficiency and robust message transmission.[2] If the task encouraged successful communication but longer words did not require more effort, learners favored longer (unambiguous) forms for both meanings regardless of their frequency; conversely, if longer words took more effort to produce but the task did not involve communication, learners favored shorter (ambiguous) forms for both meanings.

The work by Kanwal and colleagues provides an important proof of concept that production efficiency and robust message transmission can jointly shape the lexicon. Whether these results generalize to communicative systems that begin to resemble those observed in human languages is, however, an open question. The language used by Kanwal et al. (2017) contained a lexicon (participants learned two word meanings and chose between three word forms), but no grammar and none of the latent productive combinatorics that are typical in human languages. Beyond raising questions about whether the findings of Kanwal and colleagues scale up to more complex communicative systems, this work also leaves open the question of whether the encoding of grammatical relations can be affected by the balance between production efficiency and robust message transmission. The latter issue is of special interest as grammatical properties of language have often been assumed to be acquired via different mechanisms than the lexicon (Chomsky, 1981, 1986).

We address this question by studying learners' preferences in the acquisition of miniature languages that have productive combinatorics similar to (though admittedly much simpler than) natural languages. This allows us to ask, for the first time, whether the joint influences of production efficiency and robust message transmission can also shape abstract grammatical relations. The languages in the present study have productive and variable morphology and syntax, allowing the generation of hundreds (576, to be precise) of different sentences. Paralleling natural language acquisition, only a fraction of these possible meanings and forms are observed by learners, who need to *infer* the latent combinatorics (grammar) of the language.

Recent miniature language learning experiments have provided initial evidence compatible with the idea that a bias for robust message transmission influences grammatical properties of language (Fedzechkina et al., 2012, Fedzechkina, Newport, & Jaeger, 2016; Kurumada & Grimm, in press). For example, Fedzechkina, Newport, & Jaeger, 2016 exposed learners to miniature artificial languages with either fixed or flexible constituent order and optional case marking. Case marking in these languages was used with the same frequency regardless of the amount of constituent order flexibility in the language. Fedzechkina and colleagues found that learners maintained case in the flexible order language and tended to drop it in the fixed order language, mirroring the cross-linguistic correlation between case and constituent order flexibility (Blake, 2001). This finding is predicted if the utility of case marking—the balance between the effort associated with producing it and the reduction in uncertainty about sentence meaning that it results in—affects the frequency with which learners produce case. Since case marking requires effort to produce, it is more likely to be maintained when it contributes to robust message transmission by reducing uncertainty about the intended meaning (i.e., when constituent order is flexible), compared to when case marking does not contribute to uncertainty reduction (i.e., when constituent order is fixed).

However, the results of Fedzechkina and colleagues can also be explained without reference to production efficiency. In their study and similar previous work, case production always required additional effort compared to non-case-marked nouns. It is thus possible that participants were simply more likely to produce case marking when it was informative (in the flexible order language), regardless of the effort associated with case use. Thus, it remains unknown whether the cross-linguistic trade-off between constituent order flexibility and case is better explained by a *joint influence* of robust message transmission and production efficiency than by a preference for robust message transmission alone.

As this joint influence hypothesized by Zipf continues to be highly influential and untested, we seek to put it to a direct test. We manipulate *both* the amount of uncertainty about the intended message and the amount of effort required to produce case marking in a miniature artificial language. If the cross-linguistic trade-off between case and constituent order flexibility is indeed shaped by joint influences of robust message transmission and production efficiency, we would expect this trade-off to emerge only when case production requires additional effort compared to non-case-marked nouns.

## 2. Material and methods

We begin by describing participant recruitment, the miniature input grammar, the production efficiency manipulation, and the learning

[2] The task employed by Kanwal et al. (2017) intentionally confounds the *effort* and the *time* needed to produce a word. The task employed in the present work shares this property. As we review in Discussion, the same confound is observed in natural languages.

procedure.

### 2.1. Participants

Recruiting and execution of this study was approved by the Research Subjects Review Board at the University of Rochester. Participants in the experiment were self-reported native speakers of English recruited via Amazon Mechanical Turk using a custom-designed web applet (Tily, Frank, & Jaeger, 2011). Each participant was exposed to only one language and received $8 for completing the experiment, which consisted of two 40-minute sessions (for a prorated $6/hour). Following our previous work (Fedzechkina et al., 2012, Fedzechkina, Newport, & Jaeger, 2016; Fedzechkina, Chu, & Jaeger, 2017), participants were recruited until the number of successful learners reached 20 each in each of the six between-participant conditions.

Successful learning was defined exactly as in Fedzechkina, Newport, & Jaeger, 2016. This serves to reduce our researchers' degrees of freedom in determining the number of participants to recruit. The final sample submitted for analysis included the data from 120 participants across all languages and efficiency conditions (out of 143 participants who completed the experiment; see Section 2.4. for details on exclusion criteria).

### 2.2. Miniature input languages

Participants were instructed that they would be learning a novel 'alien' language by watching short videos and hearing their descriptions. The language consisted of 10 novel words (Table 1): six nouns referring to male characters (CHEF, REFEREE, BANDIT, CONDUCTOR, HUNTER, MOUNTIE), four transitive verbs (TAP, HUG, KICK, ROCK), and case markers 'di' and 'ka' (only in the *2 additional markers* condition, see below) that, if present, followed the object in the sentence. All novel words were phonotactically legal words in English, the native language of the participants. Words were synthesized individually using the Apple speech synthesizer (voice 'Alex') and concatenated into sentences in real time during the experiment. This procedure ensured that the language did not contain any prosodic cues to grammatical function assignment. All sentences in the novel language described short videos created with Poser Pro software that depicted simple transitive events such as 'hug' or 'poke' performed by two characters such as 'chef' or 'referee' (see Figs. 1 and 2 for example stimuli). All verbs occurred equally frequently with all constituent orders allowed by the input grammar and all nouns occurred equally often in the subject and object position with each verb.

Throughout the experiment, auditory descriptions were accompanied by the corresponding text in the novel language. Since the production test involved clicking on the written words available in the novel lexicon (see Section 2.3. and Fig. 1), this ensured that participants gained sufficient familiarity with novel word spellings before moving to the critical test trials.

The grammar of the language was identical to that of Fedzechkina, Newport, & Jaeger, 2016. Participants were assigned to learn one of two languages. Both languages contained optional case marking (present 67 % of the time on the object, never on the subject). As common in languages with rich case marking, the languages in the experiment had verb-final word order (Dryer & Haspelmath, 2011). The languages differed in the amount of constituent order flexibility they allowed and thus in the amount of uncertainty about the intended message based on the information provided by constituent order alone. In the flexible order language, subject-object-verb (SOV) and object-subject-verb (OSV) orders occurred with equal frequency. In the fixed order language, only SOV order was allowed. The amount of constituent order flexibility had important implications for the informativity of case marking in the two languages. In the flexible order language, constituent order was uninformative about grammatical function assignment and case marking, when present, provided important information about sentence meaning. In the fixed order language, on the other hand, constituent order was highly informative about grammatical function assignment, rendering case marking a redundant cue.

**Table 1**
The miniature lexicon.

| Nouns | Verbs | Case marker |
|---|---|---|
| barsu | skroop | di |
| forpih | velmik | ka (only in the *2 additional markers* condition) |
| doakla | kyse | |
| rizbi | tegud | |
| lanfu | | |
| peza | | |

### 2.3. Production efficiency manipulation

We further manipulated the efficiency of case production in the miniature languages—specifically the time and effort required to produce case. To do so, while holding all other properties of the languages constant, we used a novel web-based paradigm, in which participants described visual scenes by selecting words from the lexicon shown at the top of the screen (see Fig. 1). The order of lexical items was held constant on all trials within each condition—the items were always ordered alphabetically. Learners of each of the two languages were assigned to one of three efficiency conditions, resulting in a 2-by-3 between-participant design. In the *no additional effort* condition, participants were shown a case-marked and a non-case-marked variant of every noun (along with all the verbs). In this condition, production of a case-marked noun required the same number of clicks as production of a non-case-marked noun (Fig. 1A). In the *1 additional marker* condition, participants saw non-case-marked variants of all nouns along with one case marker shown as a separate word (see Fig. 1B, left). Thus, in this condition it took one additional click to produce a case-marked noun compared to a non-case-marked noun. Case marking in the *2 additional markers* condition consisted of two free markers (shown as separate words) that immediately followed each other and thus a case-marked noun required two additional clicks to produce compared to a non-case-marked noun (e.g., 'rizba di ka' vs. 'rizba', Fig. 1B, right). The representation of case markers—either as a separate word(s) or attached to the noun—was held constant across all parts of the experiment within each condition (sentence exposure, comprehension, and production tests).

If the cross-linguistic trade-off between case and constituent order flexibility is indeed shaped by the joint effect of robust message transmission and production efficiency, we expected that only learners in *1 additional marker* and *2 additional markers* conditions would use more case marking in the flexible order language (i.e., when the uncertainty about the intended sentence meaning based on the information provided by constituent order alone was high) compared to the fixed order language (i.e., when the uncertainty about the intended meaning based on the information provided by constituent order alone was low). Based on prior work within this paradigm (Hudson Kam & Newport, 2005, 2009), we expected learners in the *no additional effort* condition to match the input proportion of case marking regardless of constituent order flexibility. If, however, the cross-linguistic trade-off between constituent order flexibility and case can be fully explained by a preference for robust message transmission alone, we expected learners to use more case marking in the flexible order language compared to the fixed order language in all efficiency conditions.

### 2.4. Procedure

The miniature language learning task closely followed Fedzechkina, Newport, & Jaeger, 2016 with two exceptions: The current work used a web-based (rather than in-lab) paradigm and shortened exposure to two

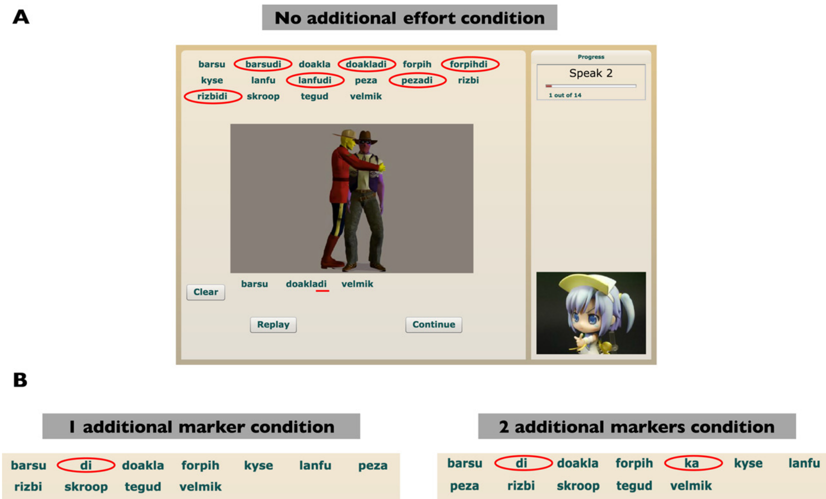


**Fig. 1.** Sentence production test in three efficiency conditions. Panel A shows a sample sentence production trial in the *no additional effort* condition. Participants saw the lexicon at the top of the screen, presented in alphabetical order (case-marked nouns are circled for illustration purposes, participants were not shown the highlighting). Sentences that participants produced via clicking on the words in the lexicon appeared on the bottom of the screen (the case marker is underlined; participants were not shown the highlighting). The image shows a still of an example video shown to participants during the sentence production test. Panel B shows how the lexicons appeared on the screen in the *1 additional marker* and *2 additional markers* conditions (case marker(s) are circled; not shown to participants).

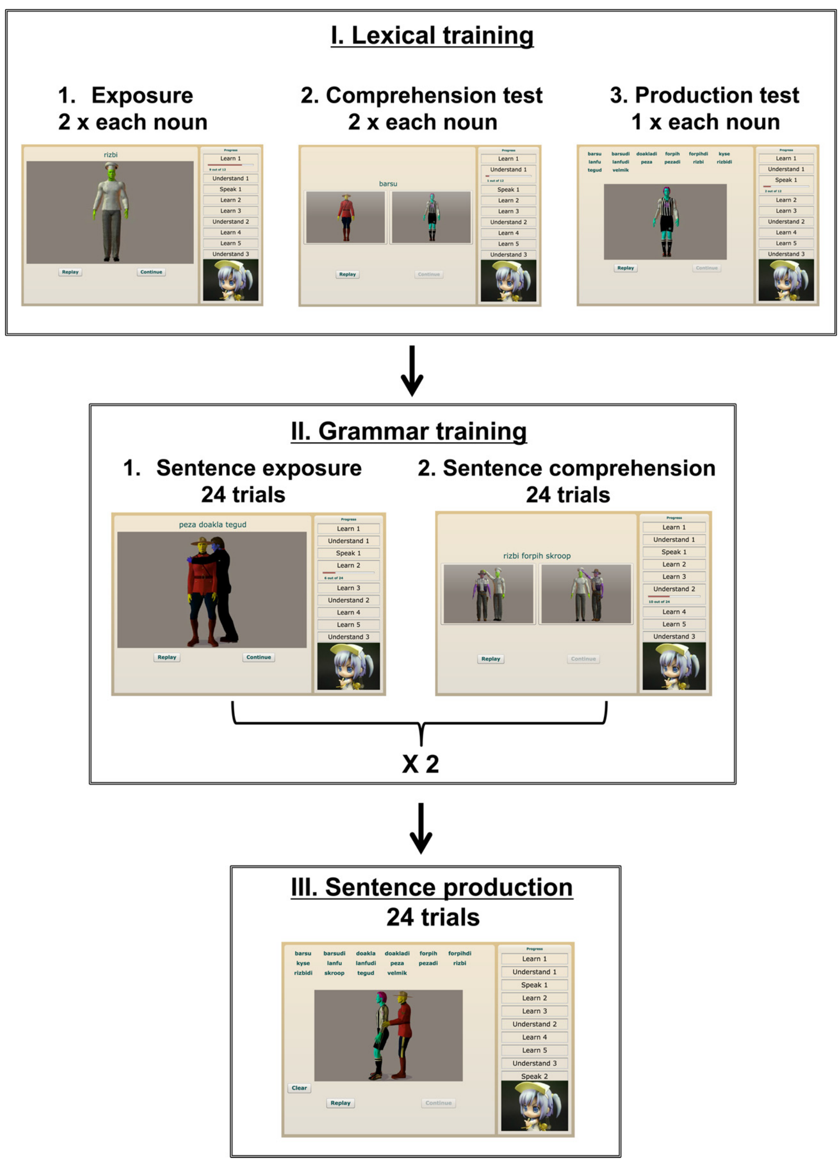


**Fig. 2.** Experimental procedure. Images represent screenshots of sample trials during each experimental block. Arrows represent the progression of experimental blocks: On both days of training, the experiment started with noun exposure and ended with sentence production.

**Table 2**
Accuracy of acquisition on the final (2nd) day of training.

| Measure | No additional effort condition | | 1 additional marker condition | | 2 additional markers condition | |
|---|---|---|---|---|---|---|
| | Fixed order | Flexible order | Fixed order | Flexible order | Fixed order | Flexible order |
| Sentence comprehension accuracy | 99 % | 97 % | 99 % | 94 % | 99 % | 91 % |
| Lexical mistakes in production | 3 % | 4 % | 1 % | 5 % | 3 % | 5 % |
| Grammatical mistakes in production | < 1 % | < 1 % | < 1 % | < 4 % | < 1 % | < 9 % |

days (compared to three) in order to increase recruitment rates. The experiment was administered in two 40-minute sessions spread over two consecutive days. The procedure was the same on both days of training (see Fig. 2 for illustration).

#### 2.4.1. Lexical training

*2.4.1.1. Noun exposure*. Participants saw pictures of characters one at a time, accompanied by their names in the novel language. Participants could listen to the names as many time as they wished before moving to the next trial.

*2.4.1.2. Noun comprehension test*. Participants were shown two pictures of characters accompanied by a name in the novel language and asked to click on the picture of the character matching the name. Participants did not receive feedback on each trial, but instead received a summary of their performance at the end of the block (e.g., 'you got 5 out of 8 right'). Participants were required to score 91 % correct on this block (no more than one incorrect trial) in order to continue to the next block. If they failed to meet the accuracy requirement, they were directed back to the noun exposure block.

*2.4.1.3. Noun production test*. Participants were shown a picture of a character with the entire lexicon of the language on the top of the screen and asked to click on the name of the character. As in the noun comprehension block, participants received a summary of feedback at the end of the block. They were required to score 91 % correct on this block in order to continue to grammar training or otherwise redirected to the noun exposure block, which they had to repeat.

#### 2.4.2. Grammar training

Participants were exposed to the novel grammar through a series of two sets of two sentence exposure blocks and one sentence comprehension block (24 trials each).

*2.4.2.1. Sentence exposure*. Participants were shown brief computer-generated videos that depicted simple transitive events accompanied by their descriptions in the novel language. Participants could replay each video (along with its description) as many times as they wanted.

*2.4.2.2. Sentence comprehension*. Participants were shown two previously unseen videos depicting the same action and participants in reversed grammatical roles. We tested comprehension on previously unseen videos to ensure that participants learned the productive combinatorics of the language. Participants heard a sentence in the novel language and were asked to click on the video that corresponded to it. No feedback at the end of the trial or feedback summary at the end of the block was provided.

*2.4.2.3. Sentence production*. Participants were shown the entire lexicon of the language at the top of the screen and asked to describe videos shown one at a time (24 in total) by clicking on the corresponding lexical items. Depending on the efficiency condition, the case marker(s) was presented either as a separate word(s) along with non-case-marked nouns or as a suffix attached to every noun along with a set of non-case-marked nouns (Fig. 1). No feedback at the end of the trial or feedback summary at the end of the block was provided. As with sentence comprehension, the videos used in the sentence production test depicted previously unseen scenes to ensure that participants learned the productive combinatorics of the language.

### 2.5. Exclusions

Participants who failed to achieve 70 % accuracy on unambiguous (i.e., case-marked) trials during the sentence comprehension test on the final day of training were removed from the analysis. This resulted in the exclusion of 22 participants (15 %), all in the flexible order language (four participants in the *no additional effort* condition, six participants *in the 1 additional marker* condition, 12 participants in the *2 additional markers* condition).

Additionally, one participant (0.6 %) in the flexible order language, *2 additional markers* condition was excluded due to experimenter error.

### 2.6. Scoring

For the remaining 120 participants, we scored their accuracy of acquisition on the final day of training to determine how well they learned the novel language. Their performance is summarized in Table 2. In sentence comprehension, we scored participants' accuracy on unambiguous trials to determine how well they learned the meaning of case marking.

In production, for each trial, we scored constituent order used in the sentence, the presence of case marking and whether it was used on the object, lexical (using incorrect vocabulary) and grammatical mistakes (using a constituent order not allowed by the input grammar or using a case marker on a constituent other than the object). All analyses were based on grammatically correct utterances only. This included sentences with lexical mistakes as long as it was possible to determine the constituent order used in the sentence. If sentence constituent order was impossible to determine (e.g., if both referent names were used incorrectly), the production was scored as both lexically and grammatically incorrect and removed from all analyses. The remaining analyses were thus conducted over sentence productions that only contained SOV or OSV orders and either no case marking or case marking on the object (95 % of all utterances of the 120 participants).

As Table 2 shows, all languages were acquired with a high degree of accuracy, suggesting that the task was feasible for our participants and that any preferences observed in case or constituent order use are unlikely to be explained by the failure to acquire the grammar of the novel language.

## 3. Results

We now turn to the analysis of learners' productions. We first analyze learners' constituent order preferences in production. We then turn to the critical prediction—whether preferences in case use in learners' productions depend on the time and effort required to produce case.

### 3.1. Constituent order in production

Our production efficiency manipulation is specific to case marker use. Thus, we predicted no effect of production efficiency on learners' constituent order preferences. Specifically we expected, following

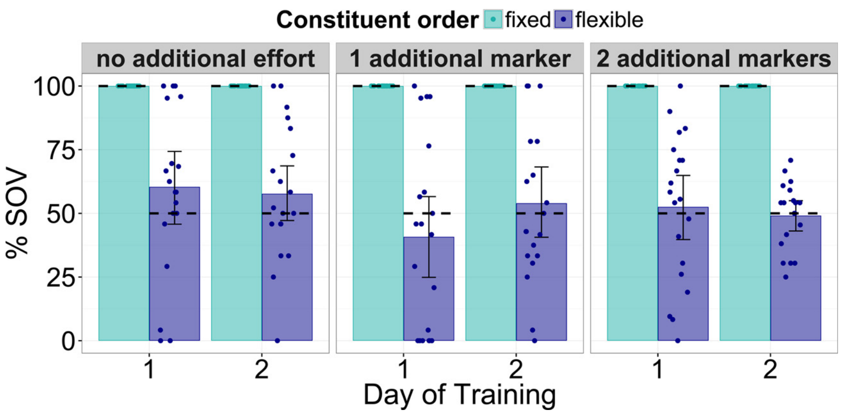


**Fig. 3.** Constituent order preferences by constituent order flexibility and efficiency condition. The dashed lines represent the input (different across languages, but same across efficiency conditions). The dots represent individual participant means. The error bars represent 95 % confidence intervals bootstrapped over by-participant means.

earlier work (Fedzechkina, Newport, & Jaeger, 2016), that learners in all efficiency conditions would match the input constituent order proportions. We test this prediction in two mutually complementing analyses. The first analysis compares learners' constituent order use across the three efficiency conditions to test whether production efficiency of case marking unexpectedly affected constituent order preferences. The second analysis directly compares learners' constituent order use to the input distribution. Both analyses are limited to the flexible order language, as learners of the fixed order language trivially matched the input SOV proportion (since only grammatical productions—i.e., only SOV—were included in the analysis, see Fig. 3).

#### 3.1.1. Constituent order use across efficiency conditions

We conducted a mixed-effects logistic regression (Breslow & Clayton, 1993; Jaeger, 2008) to predict the amount of SOV use in the flexible order language from efficiency condition, day of training, and their interaction. Efficiency condition was sliding difference coded, a contrast that compares each level of the variable to the preceding level of the variable (*1 additional marker* vs. *no additional effort* condition; *2 additional markers* vs. *1 additional marker* condition; contrast 1 = [-2, 1, 1], contrast 2 = [-1, -1, 2]). Day of training was contrast coded (1 = Day 2 vs. -1 = Day 1). The model included the maximal converging random effects structure (random intercepts by participant and by item, as well as a by-participant random slope for day and by-item random slope for day by condition interaction; items were defined as the object referents in the videos).

There was no effect of efficiency condition (*1 additional marker* vs. *no additional effort* condition: $\hat{\beta}$ = -.281, z = -1.56, p = .119; *2 additional markers* vs. *1 additional marker* condition: $\hat{\beta}$ = .112, z = .633, p = .527) or day of training ($\hat{\beta}$ = .182, z = 1.39, p = .164) on SOV use. Day of training interacted with efficiency condition (day x *1 additional marker* vs. *no additional effort* condition: $\hat{\beta}$= 246, z = 2.165, p = .03; day x *2 additional markers* vs. *1 additional marker* condition: $\hat{\beta}$ = -.218, z = -2.03, p = .042). Simple effects test revealed that on the first day of training learners in the *1 additional marker* condition used significantly less SOV order than learners in the *no additional effort* condition ($\hat{\beta}$ = -.527, z = -2.02, p = .043). This difference between efficiency conditions went away on the second day of training (all ps > 0.49).

#### 3.1.2. Constituent order compared to the input

Next, we compared SOV use by learners of the flexible order language on the final day of training to the SOV frequency in the input. We again employed mixed-effects logistic regression to predict the amount of SOV order use from efficiency condition and the maximal random effects structure (random intercept by participant and by item, as well as a by-item random slope for efficiency condition). To allow comparison against the input, efficiency condition was treatment coded and an offset term corresponding to the percentage of SOV in the input (50 %, i.e., an offset of 0 log-odds) was added to the model. Specifically, we employed three different parameterizations of the same regression, each differing only in terms of which of the three efficiency conditions was chosen as the reference level for treatment coding. In these regressions, the intercept captures whether the efficiency condition coded as reference level differs significantly from the input proportion. The analyses revealed that learners of the flexible order language matched the input proportion of SOV order in all conditions (58 % SOV production in the *no additional effort* condition: $\hat{\beta}$ = .446, z = 1.384, p = .166; 54 % SOV production in the *1 additional marker* condition: $\hat{\beta}$ = .309, z = .952, p = .341; 49 % SOV production in the *2 additional markers* condition: $\hat{\beta}$ = -.043, z = -.134, p = .893, see Fig. 3).[3]

This pattern replicates previous findings on learners' constituent order preferences (Fedzechkina, Newport, & Jaeger, 2016): Learners tend to match input constituent order distributions in all efficiency conditions. With this confirmed, we now turn to the main question of the current work—are the changes learners introduce in case distributions guided by the *joint* influences of production efficiency and robust message transmission?

### 3.2. Case marker use in production

The central prediction we set out to test is whether the cross-linguistic trade-off between case and constituent order flexibility is shaped by the *joint* influences of production efficiency and robust message transmission. In both the fixed and flexible order language, increased production effort should bias against case marking. However, in the flexible order language, unlike the fixed order language, case reduces uncertainty about the intended meaning. Under our hypothesis, we thus expect a trade-off between case and constituent order flexibility to arise in learners' productions only if case use requires additional effort compared to non-case-marked nouns. Specifically, we expect learners to produce less case when its utility is low (i.e., in the fixed order language) compared to when its utility is high (i.e., in the flexible order language) only if case production requires additional effort compared to non-case-marked nouns.

Paralleling the constituent order analyses, we test this prediction via two types of analyses. We first compare participants' case use across constituent order flexibility and efficiency conditions. We then compare learners' case use against the input they received. These analyses complement each other: Learners might have different preferences in case/constituent order use across languages and/or efficiency conditions but they may not at the same time deviate from the input they

[3] We originally presented Wilcoxon signed rank tests for comparison to the input, which return identical results. Following a reviewer's suggestions, we present the mixed-effects regressions because they take into account trial-level data, while also accounting for the repeated measures nature of the data. While an offset of 0 log-odds does not have to be explicitly specified, we use the same test below for offsets other than 0 log-odds.

receive or the other way around.

#### 3.2.1. Case marker use across languages and efficiency conditions

We first asked whether learners' preferences in case use in the fixed and flexible constituent order languages differed depending on the efficiency condition. To address this question, we ran a mixed-effects logistic regression to predict the presence of case marking from efficiency condition (Sliding difference coded, *1 additional marker* vs. *no additional effort*; *2 additional markers* vs. *1 additional marker* condition), constituent order flexibility (contrast coded, 1 = flexible vs. -1 = fixed), day of training (contrast coded, 1 = Day 2 vs. -1 = Day 1), and all interactions. The model contained the maximal converging random effects structure (random intercepts by participant and item, as well as by-participant random slopes for day and by-item random slopes for day and efficiency condition).

There was an effect of efficiency condition on case use. Learners in the *1 additional marker* condition used significantly less case overall compared to learners in the *no additional effort* condition ($\hat{\beta}$ = -.786, z = -4.255, p < .001). Learners' case use did not differ between the *2 additional markers* and *1 additional marker* conditions ($\hat{\beta}$ = .156, z = .854, p = .393). There was a main effect of constituent order flexibility on case use: Overall, learners used more case in the flexible constituent order language ($\hat{\beta}$ = .558, z = 2.5, p = .012).

Crucially, constituent order flexibility interacted with efficiency condition. The interaction was significant for the *1 additional marker* vs. *no additional effort* contrast ($\hat{\beta}$ = .437, z = 2.386, p = .017) but not for the *2 additional markers* vs. *1 additional marker* contrast ($\hat{\beta}$ = -.095, z = -.523, p = .601, see Fig. 4). Simple effects revealed that learners used significantly more case in the flexible constituent order language compared to the fixed constituent order language only in the *1 additional marker* ($\hat{\beta}$ = 1.09, z = 2.782, p = .005) and *2 additional markers* ($\hat{\beta}$ = .804, z = 2.1, p = .036) conditions, but not in the *no additional effort* condition ($\hat{\beta}$ = -.221, z = -.575, p = .565). This supports the hypothesis that the trade-off between case and constituent order flexibility emerges only if case production requires additional effort compared to non-case-marked nouns. Of note is also that the difference in case use between the fixed and flexible constituent order languages did not become more pronounced with increased effort associated with case production in the *2 additional markers* condition compared to the *1 additional marker* condition. We return to this point in the general discussion.

Additionally, there was a main effect of day of training on case use: Learners used significantly more case in their productions on the final day of training, after they became more proficient in the novel language ($\hat{\beta}$ = .296, z = 2.892, p = .004). We have observed similar learning behavior in previous work (Fedzechkina et al., 2012, Fedzechkina, Newport, & Jaeger, 2016). Day of training also interacted with efficiency condition (*1 additional marker* condition vs. *no additional effort* condition x day: $\hat{\beta}$ = .372, z = 4.256, p < .001; *2 additional markers* vs. *1 additional marker* x day: $\hat{\beta}$ = -.156, z = -1.85, p = .064). Simple effects revealed that case use increased significantly from Day 1 to Day 2 only in the *1 additional marker* ($\hat{\beta}$ = 0.823, z = 4.35, p < .001) and *2 additional markers* ($\hat{\beta}$ = .356, z = 2.09, p = .036) conditions, but not in the *no additional effort* condition ($\hat{\beta}$ = -.292, z = -1.661, p = .097). Finally, there was a marginal three-way interaction between efficiency condition, day of training and constituent order flexibility, suggesting that difference in how production efficiency affected case use in the fixed and flexible languages decreased somewhat from Day 1 to Day 2 (*1 additional marker* condition vs. *no additional effort* condition x day x constituent order flexibility: $\hat{\beta}$ = -.161, z = -1.875, p = .061). No other effects reached significance (ps > .4).

The comparison across constituent order flexibility and efficiency conditions thus supports our hypothesis: Participants produced less case in the fixed order language compared to the flexible order language, but did so only when case production required additional effort. Next, we compare case use in learners' productions to the input.

#### 3.2.2. Case marker use compared to the input

We conducted the analyses using mixed-effects regression, adding an offset term capturing the percentage of the case marking in the input (67 %). Comparisons to the input support our hypothesis. In the absence of additional effort associated with case production, learners tended to match the input. When case marking required additional effort to produce, learners tended to match the input only when the utility of case marking was high (i.e., in the flexible order language); learners tended to drop case when its utility was low (i.e. in the fixed order language).

Specifically, learners in the *no additional effort* condition matched the input proportion of case marking in their productions in both the fixed (70 % case in production, $\hat{\beta}$ = .61, z = 1.21, p = .262) and the flexible (71 % case in production, $\hat{\beta}$ = .71, z = 1.564, p = .118) order languages. This markedly differed from learners' preferences in the additional effort conditions, in which learners matched the input proportion of case marking in the flexible order language (*1 additional marker* condition: 64 % case in production, $\hat{\beta}$ = .221, z = .498, p = .619; *2 additional markers* condition: 63 % case in production, $\hat{\beta}$ = .023, z = .053, p = .958), but produced significantly less case compared to the input in the fixed order language (*1 additional marker* condition: 43 % case in production, $\hat{\beta}$ = -1.477, z = -2.979, p = .003; *2 additional markers* condition: 45 % case in production, $\hat{\beta}$ = -1.218, z = -2.493, p = .013).

In summary, the comparison of learners' case productions to the input complements the comparison across the constituent order flexibility and efficiency conditions. We find that learners generally

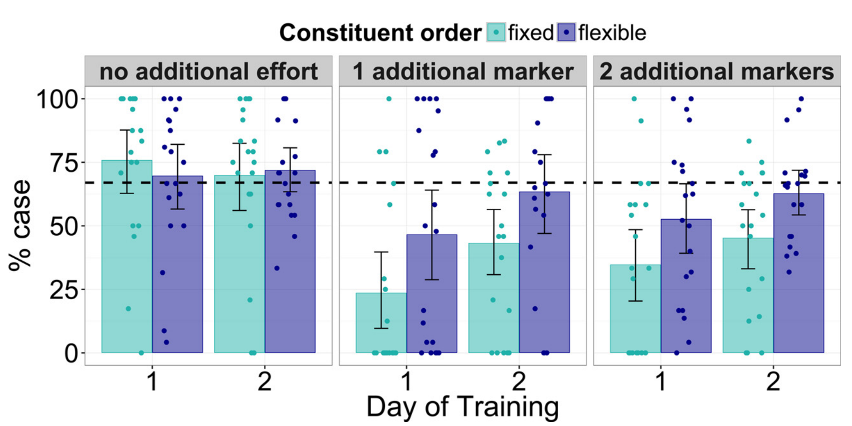


**Fig. 4.** Case marker preferences by constituent order flexibility and efficiency condition. The dashed line represents the input (same across all languages and efficiency conditions). The dots represent individual participant means. The error bars represent 95 % confidence intervals bootstrapped over by-participant means.

matched the input proportion of case marking but deviated from the input (producing less case) in conditions when the production of case required additional effort and did not reduce the uncertainty about the intended meaning (i.e., in the fixed order languages). The comparison to the input thus shows specifically how the pattern found in the first analysis comes about. Next, we discuss the likely reasons for this specific pattern before turning to the general discussion.

#### 3.2.3. Why this specific pattern?

Before we discuss the consequences of these results with regard to the joint influences of production efficiency and robust message transmission, we briefly discuss the role of a well-known third pressure in explaining the specific patterns of case use shown in Fig. 4. When case production required additional effort compared to non-case-marked nouns, learners reduced the proportion of case marking in the fixed order language. However, when case production did not require any additional effort, learners matched the input in both the flexible and fixed order languages. Why this specific pattern? One plausible explanation is found in a well-documented preference to match the distributions observed in the input (Hudson Kam & Newport, 2005, 2009) that is likely to operate in addition to the pressures for production efficiency and robust message transmission we have focused on. The preference for input matching is perhaps not surprising given that the task in miniature language learning experiments induces a goal to match the input (both in terms of grammar rules and statistics) and thus, all else being equal, 'perfect' acquisition of inconsistent patterns should result in matching the input. Similar goals are also likely to operate during natural language acquisition.

Regardless of whether the preference to match the input is specific to miniature language learning or is more general, its interaction with production efficiency and robust message transmission offers an explanation for the specific patterns of case use we observe in the present study. Specifically, this interaction can explain why learners of the fixed order languages did not reduce case to zero in the additional effort conditions. Doing so would significantly increase production efficiency but it would at the same time result in a significant deviation from the input. The same pressure to match the input could also explain why learners did not produce *more* case than the input proportion even when case marking was informative (i.e., in the flexible order languages).

Finally, a trade-off between the pressure to match the input and biases during learning might explain why learners in the *2 additional markers* condition do not express the balance between robust message transmission and efficiency more strongly compared to the *1 additional marker* condition, in which case production required less effort (one fewer click). Recall, that in the *2 additional markers* condition, the free markers were always displayed in the fixed position on the screen and that the path between the two case markers (i.e., the path from 'di' to 'ka') in the visual display was short (spanning two words, see Fig. 1B). Thus, it is possible that since it was the same targeting problem across trials, participants learned a motor routine for the production of the second marker in the *2 additional markers* condition. One could not as easily learn a motor routine for the first marker in the *1 additional marker* and *2 additional marker* conditions: While this marker was always in the same position on the screen, its motor planning started from different locations on the screen as different nouns needed to be case-marked on different trials. It is thus possible that the amount of additional effort required to produce an additional case marker in the *2 additional markers* condition was not sufficient to warrant additional deviations from the input beyond the *1 additional marker* condition.

## 4. Discussion

Learners' preferences in case use support the hypothesis that the correlation between constituent order flexibility and case marking is caused by the *joint* influence of production efficiency and robust message transmission, rather than by one of the two pressures alone. The trade-off between case and constituent order flexibility emerged in learners' productions only when case use required additional effort compared to the production of non-case-marked nouns. When case required no additional effort compared to non-case-marked nouns, learners matched the input proportion of case. The present findings thus follow—at least for the example studied here—Zipf's critical prediction that preferences for production efficiency and robust message transmission jointly give rise to cross-linguistically observed grammatical patterns.

This adds to the growing body of work suggesting that the preference to balance robust message transmission and production efficiency plays a role in shaping properties of natural language grammars. The present findings replicate and extend the work by Fedzechkina, Newport, & Jaeger, 2016, who found that learners tended to drop case if the language had fixed order but maintained case if the language had flexible order. The present work also mirrors and extends the findings of Kanwal et al. (2017), who found that word frequency traded off against word length in a small (two words) artificial lexicon. We find a similar trade-off at work in more complex miniature languages that include a larger lexicon and morpho-syntactic contingencies. This trade-off is observed during sentence production and therefore goes beyond the selection between different candidate forms for an isolated word. Thus, our results suggest that the joint effect of production efficiency and robust message transmission goes beyond the lexicon, also affecting morpho-syntactic grammatical preferences.

In Appendix A in Supplementary materials, we present evidence that not all production patterns that, at first blush, seem *consistent* with a joint influence of production efficiency and robust message transmission necessarily involve both pressures. Following observations made in Fedzechkina, Newport, & Jaeger, 2016, we analyze learners' preferences for the use of case *conditional* on sentence constituent order. Replicating our previous work, we find that learners of the flexible order language tend to condition case marking on constituent order and use more case in OSV compared to SOV sentences. This preference is independent of—holds over and above—the results of our main analysis of case use. We find that this preference for asymmetric case systems does *not* depend on our manipulation of production efficiency (unlike learners' preference to trade off constituent order flexibility and case when it is associated with additional production effort). This validates the premise of the present study.

In the remainder, we first discuss potential concerns with our interpretation of the results (some raised by reviewers). We then revisit our operationalization of production efficiency and discuss how future work might distinguish between various factors that contribute to it. Finally, we outline how future work might expand the paradigm we have employed here to further explore the influences of robust message transmission and production efficiency.

### 4.1. Did learning difficulty confound our results?

One potential concern is that the patterns of case use we observed in learners' productions might have been confounded by the relative difficulty of the fixed and flexible order languages. Recall that participants who scored less than 70 % correct on case-marked trials on the sentence comprehension test on the final day of training were removed from the analysis. All of these participants (22 across the three efficiency conditions) were learners of the flexible order language, suggesting that this language was harder to learn than the fixed order language. This itself is not surprising—our participants are native speakers of English (a fixed order language with no case marking). Could the relative ease of the fixed order language imply that our exclusion criteria did not exclude bad learners for this language, whereas we successfully filtered out bad learners of the flexible order language? If better learners, who are more likely to learn the function of the case marker, are also more likely to *use* the case marker, this would explain why we observed more case marking in the flexible order language. Critically though, this would *not* explain the finding of primary interest—the fact that constituent order flexibility *interacts* with production efficiency in predicting case use (see Fig. 4): While learners of the flexible order language produced the same the amount of case marking across all efficiency conditions, learners of the fixed order language tended to drop case once production effort became a consideration (i.e., in the additional effort conditions).

At the suggestion of an anonymous reviewer, we further assessed whether the observed interaction between constituent order flexibility and production efficiency can be explained by the accuracy of learning. We repeated the analyses of case marker use reported in Section 3.2.1 while including each participant's comprehension accuracy (as a proxy for learning accuracy) as a control predictor. Consistent with the idea that better learners are more likely to produce case marking, comprehension accuracy was a significant positive predictor of case use. However, all findings reported above—including the critical interaction between constituent order flexibility and production efficiency—remained statistically significant and similar in effect size to the original model (see Appendix C for details in Supplementary materials). This suggests that our conclusions are not confounded by learning success.

A second potential concern is that our manipulation of production efficiency resulted in visual displays that differed across the efficiency conditions. In the *no additional effort* condition, the case marker was attached to the noun, while in the additional effort conditions, the case marker was presented as a separate word. It is plausible that the visual presentation of the case marker (separate from the stem or not) might have affected learning. For example, the separate presentation of the case marker(s) in the *1 additional marker* and *2 additional markers* conditions might have helped participants learn that some forms in the input (the case markers) followed certain morpho-syntactic rules. This would not, however, explain our critical result that learners' use of case marking in the additional effort conditions depended only on the amount of constituent order flexibility. If anything, better learning of case marking should have resulted in more probability matching in the additional effort conditions. Similarly, hypothetical failures to find or remember to produce the case marker in the additional effort conditions do not explain why case use depended on both constituent order flexibility and production efficiency.

### 4.2. What is production efficiency?

Our manipulation of production efficiency does not clearly distinguish between motor effort, cognitive effort, and production speed. Case production in the additional effort conditions was associated with more mouse movements and additional clicking, resulting in additional motor effort. Case production in these conditions was, however, also associated with additional visual search effort as participants had to either find the case markers in the visual display or remember where on the display the case markers are located. While this additional cognitive effort was arguably minimal—recall that the case markers appeared in the same position on every display (Fig. 1B)—it is theoretically possible that it contributed to our results. Regardless of whether additional visual search or memory actually played a direct role in the perceived effort of case use, they likely contributed to the fact that case production resulted in additional *time* compared to production of non-case-marked nouns. Our paradigm therefore does not distinguish between the effort and speed required for case production (and thus also the speed of communication).

The present work shares this property with the most closely related previous work. In the study by Kanwal et al. (2017), participants had to select words by clicking on them. In the conditions that emphasized efficiency, words were 'sent' (communicated) by clicking on the word and holding the mouse button. The word would then appear letter by letter and the next trial only started when the full word had appeared. It took 8.4 s to send a long word form (e.g., *zopudon*)—more than double the time required to send a short form (e.g., *zop*, 3.6 s). While this manipulation emphasized speed it also entailed more effort (holding down the button for a longer time).

Both the present paradigm and the paradigm of Kanwal and colleagues thus confound effort and speed, though arguably the two paradigms create different emphases. The manipulation of time required to communicate was much shorter in our study compared to that of Kanwal and colleagues. Specifically, it is quite unlikely that the time required to produce a case-marked noun was twice as long as the time required for a non-case-marked noun in our set up. This was supported by a post-hoc analysis, which found no evidence that visual search time affected case production (see Appendix B in Supplementary materials).

Critically, a confound between effort and speed is also often present in natural case use and many similar signal reduction phenomena in human languages [for a review, discussion, and exceptions, see Jaeger and Buz (2017)]. Indeed, time itself can be seen as a resource the use of which cognitive systems try to optimize (Griffiths, Lieder, & Goodman, 2015) and balance against (in this case, the goal of communication would be balanced against time). Both the paradigm by Kanwal and colleagues and the present paradigm offer rich possibilities for future work aimed at disentangling the effects of effort and time.

Similarly, other miniature artificial language paradigms offer promising avenues for disentangling the effects of motor effort (clicking) and cognitive effort (visual search) in our experiment. One such paradigm is that of Harmon and Kapatsinski (2017) used to study frequency-based semantic extension. They find specifically that learners of a novel miniature language prefer to re-use a frequent form in the input (e.g., a singular diminutive suffix) over an infrequent one (e.g., a plural non-diminutive suffix) to mark a previously unseen meaning (e.g., diminutive plural). Importantly, they find that this frequency-based semantic extension only emerges if frequency is associated with greater form accessibility (and thus, with lower effort required to retrieve the form): Harmon and Kapatsinski (2017) observed this effect in production, when learners needed to retrieve the forms from memory, but not in a 2-alternative forced-choice task, when both forms were presented to the learners. Similar logic can be applied to the present task: By highlighting or making case markers otherwise more available, we can begin to meaningfully tease apart the effects of cognitive and motor effort on case use.

### 4.3. Future directions in studying the influences of communicative efficiency on language structure

Our hypothesis—that the cross-linguistic trade-off between case and constituent order flexibility is caused by an efficient balance between robust message transmission and production efficiency—is motivated by theories of efficient communication (e.g., Shannon, 1948; Zipf, 1949). However, the task we employed to test this hypothesis did not involve explicit communication. In this respect, the current study is different from similar work on the lexicon structure by Kanwal et al. (2017), where two participants communicated with each other by transmitting picture descriptions.

This is not to say that no communicative pressures were present in our paradigm: Speakers design their utterances to take into account not only their direct communicative partner, but also potential over-hearers such as the experimenter or bystanders who are not immediately involved in the conversation, provided that there is a goal to communicate (Clark & Murphy, 1982; Clark & Schaefer, 1992). Our task was designed to introduce a clear goal to produce communicative utterances – participants were instructed to talk to an alien informant. Our findings suggest that we were successful in this – learners in our experiment produced case marking in a way that is predicted by communicative efficiency accounts. Nevertheless, our task under-emphasized communicative pressures compared to natural language use. Thus, paradigms like the one employed by Kanwal and colleagues are predicted to find as strong or stronger effects of communicative pressures for input languages similar to ours.

## 5. Conclusions

Our findings add to the growing body of literature suggesting that some cross-linguistic properties originate in the limitations of human cognitive systems such as preferences for efficient processing and communication (Fedzechkina et al., 2012, Fedzechkina, Newport, & Jaeger, 2016; Fedzechkina, Chu, & Jaeger, 2017; Fenk-Oczlon, 2001; Ferrer i Cancho, 2005; Futrell, Mahowald, & Gibson, 2015; Gibson et al., 2013; Jäger, 2007; Kurumada & Jaeger, 2015; Maurits et al., 2010). We also show how by using miniature language learning methodology we can begin to meaningfully tease apart competing hypotheses about the same phenomenon. The formalization of efficiency in our experiment in terms of additional clicks is certainly an imperfect approximation of production effort in terms of additional phonemes or syllables. It allowed us, however, to model scenarios, in which case production either involved additional effort or not, which would be difficult to tease apart in natural language production (there is always additional effort associated with articulating an additional phoneme or syllable). By finding that the cross-linguistic trade-off between case and constituent order flexibility emerges only when additional effort is associated with case production, the present findings provide support for the idea that some cross-linguistic properties of morpho-syntax are better explained by a balance between production efficiency and robust message transmission rather than by accounts that do not require a reference to production efficiency (e.g., Bates & MacWhinney, 1989). Our work, thus, provides further support for the suitability of miniature language learning paradigms to test proposals that have been challenging to evaluate on natural language data.

### Declaration of Competing Interest

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

### Acknowledgement

This work was supported in part by NSF CAREER grant IIS-1150028 to T. Florian Jaeger.

### Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.cognition.2019.104115.

### References

Aissen, J. (2003). Differential Object Marking: Iconicity vs. Economy. *Natural Language and Linguistic Theory, 21*(3), 435–483. https://doi.org/10.1023/A:1024109008573.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner, & L. Gleitman (Eds.). *Language acquisition: The state of the art* (pp. 173–218). Cambridge: Cambridge University Press.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. *The crosslinguistic study of sentence processing,* 3–73.

Blake, B. J. (2001). *Case.* Cambridge: CUP.

Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association, 88*, 9–25. https://doi.org/10.1080/01621459.1993.10594284.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language.* New York: Praeger.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In J. F. LeNy, & W. Kintsch (Eds.). *Language and comprehension* (pp. 287–299). Amsterdam: North Holland Publishing.

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

Cohen Priva, U. (2015). Informativity affects consonant duration and deletion rates. *Laboratory Phonology, 6*(2), 243–278.

Croft, W. (2003). *Typology and universals.* Cambridge: CUP.

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

Currie-Hall, K., Hume, E., Jaeger, T. F., & Wedel, A. (2018). *The role of predictability in shaping phonological patterns.* Linguistic Vanguard.

Dryer, M., & Haspelmath, M. (Eds.). (2011). *The world atlas of language structures online*. Munich: Max Planck Digital Library.

Fedzechkina, M., Chu, B., & Jaeger, T. F. (2017). Human information processing shapes language change. *Psychological Science, 29*(1), 72–82.

Fedzechkina, M., Jaeger, T. F., & Newport, E. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences of the United States of America, 109*(44), 17897–17902. https://doi.org/10.1073/pnas.1215776109.

Fedzechkina, M., Newport, E., & Jaeger, T. F. (2016). Balancing effort and information transmission during language acquisition: Evidence from word order and case-marking. *Cognitive Science, 41*(2), 416–446. https://doi.org/10.1111/cogs.12346.

Fenk-Oczlon, G. (2001). *Familiarity, information flow, and linguistic form. Frequency and the emergence of linguistic structure*431–448.

Ferrer i Cancho, R. (2005). Zipf's law from a communicative phase transition. *European Physical Journal B: Condensed Matter and Complex Systems, 47*(3), 449–457.

Ferrer i Cancho, R., & Moscoso del Prado Martin, F. (2013). Information content versus word length in random typing. *The Journal of Statistical Mechanics: Theory and Experiment, L12002*. https://doi.org/10.1088/1742-5468/2011/12.

Finley, S., & Badecker, W. (2008). Substantive biases for vowel harmony languages. In N. Abner, & J. Bishop (Eds.). *27th West Coast conference on formal linguistics* (pp. 168–176). Somerville, MA: Cascadilla Press.

Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences of the United States of America, 112*(33), 10336–10341.

Gabelentz, v.d (1901). *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherige Ergebnisse.* Leipzig: C. H. Tauchnitz.

Gibson, E., Piantadosi, S., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of cross-linguistic word order variation. *Psychological Science, 4*(7), 1079–1088. https://doi.org/10.1177/0956797612463705.

Givón, T. (1985). Iconicity, isomorphism and non-arbitrary coding in syntaz. In J. Haiman (Ed.). *Iconicity in syntax* (pp. 187–219). Amsterdam: John Benjamins.

Gomez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109–135.

Griffiths, T., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science, 7*(2), 217–229.

Harmon, Z., & Kapatsinski, V. (2017). Putting old tools to novel uses: The role of form accessibility in semantic extension. *Cognitive Psychology, 98*, 22–44.

Hawkins, J. A. (2004). *Efficiency and complexity in grammar.* Oxford: Oxford University Press.

Hockett, C. (1958). *A course in modern linguistics.* New York: MacMillan.

Hudson Kam, C., & Newport, E. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development, 1*(2), 151–195. https://doi.org/10.1080/15475441.2005.9684215.

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

Hupp, J. M., Sloutsky, V. M., & Culicover, P. W. (2009). Evidence for a domain-general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes, 24*(6), 876–909. https://doi.org/10.1080/01690960902719267.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446. https://doi.org/10.1016/j.jml.2007.11.007.

Jaeger, T. F., & Buz, E. (2017). Signal reduction and linguistic encoding. In E. M. Fernandez, & H. S. Cairns (Eds.). *Handbook of psycholinguistics*. Wiley-Blackwell.

Jäger, G. (2007). Evolutionary Game Theory and Typology. A Case Study. *Language, 83*(1), 74–109. https://doi.org/10.1353/lan.2007.0020.

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., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition, 141*, 87–102.

Kurumada, C., & Grimm, S. (2019). Predictability of meaning in grammatical encoding: Optional plural marking. *Cognition, 136*(2015), 215–221.

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

Manin, D. (2006). Experiments on predictability of word in context and information rate in natural language. *Information Processes, 6*(3), 229–236.

Maurits, L., Perfors, A., & Navarro, D. (2010). Why are some word orders more common than others? A uniform information density account. In J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, & A. Culotta (Vol. Eds.), *Advances in neural information processing systems: 23*, (pp. 1585–1593). Vancouver, BC.

Meylan, S., & Griffiths, T. (2017). Word forms-not just their lengths-are optimized for efficient communication. *ArXiv, 1703*, 01694. e-prints, arXiv.

Piantadosi, S. (2014). Zipf's word frequency law in natural language: A critical review and future directions. *Psychological Bulletin & Review, 21*, 1112–1130.

Piantadosi, S., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences of the United States of America, 108*(9), 3526. https://doi.org/10.1073/pnas.1012551108.

Piantadosi, S., Tily, H., & Gibson, E. (2013). Information content versus word length in natural language: A reply to Ferrer-i-Cancho and Moscoso del Prado Martin. *ArXiv, arXiv:1307.6726* e-prints.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Shannon, C. (1948). A mathematical theory of communications. *Bell Labs Technical Journal, 27*(4), 623–656.
Slobin, D. (1977). Language change in childhood and in history. In J. Macnamara (Ed.). *Language learning and thought* (pp. 185–214). New York: Acad. Press.
Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition, 116*(3), 444–449. https://doi.org/10.1016/j.cognition.2010.06.004.
Soskuthy, M., & Hay, J. (2017). Changing word usage predicts changing word durations in New Zealand English. *Cognition, 166*, 298–313.
St Clair, M. C., Monaghan, P., & Ramscar, M. (2009). Relationships between language structure and language learning: The suffixing preference and grammatical categorization. *Cognitive Science, 33*(7), 1317–1329. https://doi.org/10.1111/j.1551-6709.2009.01065.x.
Tily, H., Frank, M., & Jaeger, T. F. (2011). The learnability of constructed languages reflects typological patterns. In L. Carlson, C. Hoelscher, & T. Shipley (Eds.). *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 1364–1369). .
Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition, 128*, 179–186.
Zipf, G. K. (1949). *Human behavior and the principle of least effort.* Cambridge, MA: Addison-Wesley.