



Redundancy can benefit learning: Evidence from word order and case marking

Shira Tal^{a,b,*}, Inbal Arnon^c

^a Department of Cognitive Sciences, The Hebrew University of Jerusalem, Israel

^b Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, United Kingdom

^c Department of Psychology, The Hebrew University of Jerusalem, Israel

ARTICLE INFO

Keywords:

Redundancy
Artificial language learning
Language acquisition
Learning biases

ABSTRACT

The prevalence of redundancy in the world languages has long puzzled language researchers. It is especially surprising in light of the growing evidence on speakers' tendency to avoid redundant elements in production (omitting or reducing more predictable elements). Here, we propose that redundancy can be functional for learning. In particular, we argue that redundant cues can facilitate learning, even when they make the language system more complicated. This prediction is further motivated by the Linguistic Niche Hypothesis (Lupyan & Dale, 2010), which suggests that morphological complexity can arise due to the advantage redundancy might confer for child learners. We test these hypotheses in an artificial language learning study with children and adults, where either word order alone or both word order and case marking serve as cues for thematic assignment in a novel construction. We predict, and find, that children learning the redundant language learn to produce it, and show better comprehension of the novel thematic assignment than children learning the non-redundant language, despite having to learn an additional morpheme. Children in both conditions were similarly accurate in producing the novel word order, suggesting redundancy might have a differential effect on comprehension and production. Adults did not show better learning in the redundant condition, most likely because they were at ceiling in both conditions. We discuss implications for theories of language learning and language change.

1. Introduction

Languages are claimed to be shaped for efficient use, striking an optimal balance between speakers' preference to minimize production effort, and their need to be understood by their interlocutors (Gibson et al., 2019; Givón, 1991; Haspelmath, 2021; Jaeger & Buz, 2017; Mahowald, Dautriche, Braginsky, & Gibson, 2020; Zipf, 1949). An efficient way of balancing between these two opposing pressures is to assign less linguistic material to more predictable meanings, thereby putting less effort into less challenging messages (for reviews see Gibson et al., 2019; Jaeger & Buz, 2017). In information-theoretic terms, that means speakers should aim to minimize redundancy in their productions: They should omit or reduce more predictable linguistic material when they can do so without changing the message (Caballero & Kapatsinski, 2015; Dahl, 2004; Gibson et al., 2019; Jaeger, 2010; Pate & Goldwater, 2015). Indeed, there is ample evidence suggesting that speakers adhere to this principle across different levels of linguistic

analysis (Aylett & Turk, 2004; Cohen Priva, 2015; Frank & Jaeger, 2008; Jaeger, 2010; Kravtchenko, 2014; Mahowald, Fedorenko, Piantadosi, & Gibson, 2013). For example, English speakers are more likely to contract morphological elements (isn't vs. is not) when the meaning of the contractible element is more predictable in context (Frank & Jaeger, 2008).

At the same time, however, redundancy is attested in different forms across languages (Hengeveld & Leufkens, 2018; Leufkens, 2020; Trudgill, 2009; Wit & Gillette, 1999). For example, many languages have agreement patterns between an argument and a verb even when both are obligatory and mark the same meaning (Haig & Forker, 2018). English, for instance, requires the morpheme *-s* to mark third person singular on verbs (e.g., *the man walks*), even though both features (third person and singularity) can be deduced from the subject, whose expression is obligatory. Another example that is attested in a number of typologically diverse languages is a phenomenon known as *multiple exponence*, where a single meaning is redundantly marked using

* Corresponding author at: 3 Charles Street, Edinburgh EH8 9AD, United Kingdom.

E-mail address: stal@ed.ac.uk (S. Tal).

multiple morphological markers (Caballero & Harris, 2012; Caballero & Kapatsinski, 2015; Harris, 2017). In Choguita Rarámuri, for example, words containing an inner derivational marker for causatives and applicatives can have a second, optional, marker suffixed to the noun (e.g., 'su-n-ma' vs. 'su-n-ki-ma', both meaning 's/he will sew it for X'. Caballero & Kapatsinski, 2015). Importantly, languages can also redundantly mark grammatical information: One such example is the encoding of thematic assignment (who is doing what to whom) by both word order and case marking (Levshina, 2020; Siewierska, 1998). Japanese, for instance, has case marking on top of a relatively rigid word order (Levshina, 2020).

How can the presence of redundancy in language structure be reconciled with speakers' tendency to avoid it in production? The answer to this question may lie in understanding the function of redundancy in language. Here, we propose that redundancy can be functional for learning. In particular, we suggest that (a) speakers increase (or maintain) the use of redundant cues when conversing with learners, and (b) redundancy can facilitate learning under certain conditions. This proposal is compatible with the principles of efficient communication: The balance between effort and understandability should change depending on the comprehension ease or difficulty within a conversation (Aylett & Turk, 2004; Buz, Tanenhaus, & Jaeger, 2016; Kurumada & Jaeger, 2015). In particular, speakers should use more linguistic material when their listener is seen as having more difficulty in comprehension, as in the case of learners. The first prediction has gained support from several recent studies, showing that speech directed to learners is more redundant than speech directed to proficient speakers (Pate & Goldwater, 2015; Tal, Grossman, & Arnon, 2021; Tal, Grossman, Rohde, & Arnon, 2021). There is also evidence that redundant cues can facilitate learning across several linguistic and non-linguistic domains. However, there is less evidence for the facilitative effect of redundant *linguistic* cues, where the cues themselves have to be used as part of the linguistic system. In the next section we review this literature in detail and outline our research question.

1.1. Redundancy and language learning

The advantage of multiple cues in learning has been demonstrated for different cognitive domains such as vision (Sloutsky & Robinson, 2013) and category formation (Yoshida & Smith, 2005). There is also evidence suggesting that redundant cues benefit various aspects of language learning. First, various studies document the facilitative role of multimodal cues for language learning (Murgiano, Motamedi, & Vigliocco, 2020, for a related discussion on the role of multimodal cues in language processing, see Holler & Levinson, 2019). In particular, it has been suggested that having coordinated presentation of the same information across more than one modality facilitates infants' attention and learning (The Intersensory Redundancy Hypothesis, Bahrack & Lickliter, 2000; Bahrack, Lickliter, & Flom, 2004; Flom & Bahrack, 2010). For example, melody facilitates learning of lyrics (Thiessen & Saffran, 2009) and tactile cues facilitate statistical learning of tone co-occurrence patterns (Lew-Williams, Ferguson, Abu-Zhaya, & Seidl, 2019). A recent computational model demonstrated that multimodal cues can benefit mappings between forms and meanings (Monaghan, 2017). The cues were probabilistic in the learning phase (appeared only some of the time), and were absent during testing (where only the labels appeared). In line with our prediction, the computational model showed that words were learned better from multiple cues (pointing, prosody, and distributional cues) compared to single cues.

Less work has asked whether redundant *linguistic* cues (presented in the same modality) are also advantageous for learning, though this has been suggested (Audring, 2014; Bates & MacWhinney, 1989; Christiansen & Monaghan, 2006), and there is some evidence to support this idea. Morgan, Meier, and Newport (1987) had adult participants learn phrase-structure rules in an artificial language. For some of the participants the training input (but not the test) contained additional

morphological or prosodic cues to phrase boundary. They found that when participants' input included redundant cues to phrase structure, they learned it successfully, but did not manage to learn it when there was no redundancy. A similar pattern was reported by Gerken, Wilson, and Lewis (2005) in work with infants: 17-month-old infants were only able to assign nouns to the correct gender class when gender was marked redundantly by both a word ending and a suffix.

While these examples support the general notion that redundancy benefits language learning, they differ from our research question in an important way. The multiple cues used in these studies did not have to be learned in and of themselves: Learners did not have to reproduce them themselves (e.g., even if a word was learned together with a pointing gesture, the learner was not expected to reproduce the pointing together with the new meaning). In these studies, we cannot tell whether the redundant cue itself was fully mastered (i.e. used correctly in production and perception), and whether mastering it facilitated learning of the whole system (see discussion in Leufkens, 2020). In natural languages, redundant linguistic cues need to be learned as part of the linguistic system. When learning a language with a redundant case marking cue, for example, case marking has to be learned as part of the language. Unlike a pointing gesture which can direct attention to a certain element even if pointing is not used in subsequent productions, the learner is expected to learn both how to use case marking to identify thematic assignment, and how to produce the correct case marking when they speak. For this reason, languages with redundant morphemes are considered more complex (Bentz, Ruzsics, Kopleinig, & Samardzi, 2016; Dahl, 2004; Lupyan & Dale, 2010), and the possible learning advantage they confer is debatable (Leufkens, 2020). On the one hand, redundancy affords the listener additional cues for the relation to be learned (Audring, 2014; Bates & MacWhinney, 1989; MacWhinney, 1987). On the other hand, the addition of redundant cues adds complexity and can make the system less transparent, for instance, by hampering systematic one-to-one mapping between meaning and form (Hengeveld & Leufkens, 2018; Leufkens, 2020). Leufkens (2020) suggests that the seemingly contradictory effect redundancy may have on learning can be partly resolved by differentiating two aspects of learning. While learning the *redundantly marked feature* can be facilitated by redundancy (e.g., the acquisition of grammatical gender), learning the *redundant cue itself* might not be (e.g., the acquisition of the rules for redundant gender marking). A similar distinction between the advantage of cues for learning as opposed to the learnability of the cues themselves is made in the Competition Model (E. Bates & MacWhinney, 1989; MacWhinney, 1987; MacWhinney, 2001). In this framework, language comprehension is based on the detection of a series of cues. Comprehension is impacted by the reliability and availability of the cues at hand, and the way they interact with one another. However, at the same time, cues also have a *cost*: They themselves have to be perceived, processed and stored in memory (E. Bates & MacWhinney, 1989). If a cue is very costly, it can be abandoned despite its informativity.

Given the distinction between learning something a redundant cue directs attention to, and learning the redundant cues themselves, we go beyond previous work to ask whether redundant linguistic cues can facilitate language learning despite the added cost of having to learn and produce an additional element. We predict that learners will produce a redundant cue (indicating they learned it), and that such redundancy will facilitate comprehension, despite the added complexity of learning another cue (see discussion of possible differences between comprehension and production in Section 1.4 below). In line with previous work, we adopt the information-theoretic notion of redundancy where omission and reduction of linguistic material (here specifically, morphemes) with no change to the message is seen as reducing redundancy (Caballero & Kapatsinski, 2015; Frank & Jaeger, 2008; Jaeger, 2010; Kurumada & Jaeger, 2015). Before presenting our case study in detail, we turn to discuss a further motivation for testing the relation between redundancy and learning.

1.2. Are all learners alike? – The linguistic niche hypothesis

Examining the link between redundancy and learning is also motivated by a recent influential proposal about the impact of different kinds of learners on the morphological complexity of a language. A growing body of work suggests that languages spoken by a larger proportion of adult learners seem to become more morphologically simplified over time (Bentz, Verkerk, Kiela, Hill, & Buttery, 2015; Bentz & Winter, 2013; Dale & Lupyan, 2012; Lupyan & Dale, 2010; Trudgill, 2009, 2011; Wray & Grace, 2007). The Linguistic Niche Hypothesis (LNH, Dale & Lupyan, 2012; Lupyan & Dale, 2010) accounts for this tendency and offers a possible solution for an additional important question: Why are some languages more morphologically complex than others? The LNH accounts for this by proposing a causal link between the prominent type of language learners in a community and the degree of morphological complexity of the language. The hypothesis is that languages adapt themselves over time to serve the particular needs of their learners. According to this account, children and adults differ in the impact of redundancy on learning: While child learners benefit from redundancy, adult learners do not (or at least not to the same degree). As a result, languages spoken by larger proportion of adult learners are predicted to lose morphological complexity over time, while languages spoken by a smaller proportion of adult learners (hence, more child learners) are predicted to sustain it. While intuitively appealing, there is no direct evidence that children benefit from redundant cues in learning whereas adults do not. Here, we will test the impact of redundancy on both child and adult learners, to see (a) whether redundancy benefits learning, and (b) whether it is facilitative to a different degree for child and adult learners. We focus on the morpho-syntactic redundancy of word order and case marking and use the transitive construction as a test case.

1.3. The transitive construction

Languages use different cues to indicate who-did-what-to-whom in the transitive construction (Levshina, 2020). Two prominent cues are word order and case marking (Ibbotson & Tomasello, 2009; Levshina, 2020). Investigating learning when both cues are used as opposed to only one of them allows us to examine the possible advantage of redundant marking. The contribution of different cues to sentence interpretation has been studied extensively within The Competition Model (Bates, McNew, MacWhinney, Devescovi, & Smith, 1982; MacWhinney, 1987; MacWhinney, 2001). In this framework, comprehension of thematic assignment is based on the detection of cues identifying the agent and the patient, like case marking and word order. Comprehension is determined by the reliability and availability of these cues, and whether the cues work in coalition (mark the same meaning) or competition. Coalition of cues is predicted to facilitate comprehension of thematic assignment (Bates et al., 1982; Bates & MacWhinney, 1989; Ibbotson & Tomasello, 2009). Stemming from this theoretical framework, studies in various languages have tested how children utilize cues to comprehend transitive constructions. Dittmar, Abbot-Smith, Lieven, and Tomasello (2008) examined the relative reliance of German-speaking toddlers on word order and case marking. They found that 2;6-year-olds could comprehend transitive sentences only when both cues were used. The findings that children tend to rely on coalition of cues were replicated across several languages (Cantonese: Chan, Lieven, & Tomasello, 2009; Japanese: Matsuo, Kita, Shinya, Wood, & Naigles, 2012; Warlpiri: O'Shannessy, 2010). These findings seem to support the hypothesis that redundant cues benefit learning. However, in many of these cases, the redundant form is also the prototypical and the most frequent form in child-directed speech (Dittmar et al., 2008; Ibbotson & Tomasello, 2009). It is therefore not clear whether comprehension was

facilitated because of the redundant cues, or because of the greater familiarity of young children with these constructions (which happened to also have redundant cues). In the current study we tease these two factors apart by conducting an artificial language learning experiment.

1.4. The current study

In the current study, we use an artificial language learning task to assess the impact of redundant morpho-syntactic cues on children and adults learning a novel language. We use the same task with both age groups to enable direct comparison between them. For each age group we compare the learnability of transitive constructions in two versions of an artificial language with fixed OSV word order. In one version, there were no additional cues to thematic assignment (the control language), and in the other, there was an additional case marking cue on objects (the redundant language): This cue was redundant because the fixed word order means that thematic assignment could be fully deduced from the linear ordering of the words. We use OSV word order because it differs from the dominant word order of Hebrew (SVO) — the language of the participants in our study. Following exposure, we test both comprehension and production of sentences in the novel language. Our main prediction is that the presence of a redundant cue will facilitate comprehension of thematic assignment: We predict that learners of the redundant language will use the case marking cue in production (indicating they learned it as part of the linguistic system), and more importantly, that they will show better comprehension of the thematic assignment compared to learners of the control language. The redundant language should be advantageous even though the control language is simpler by not containing an additional cue to perceive or learn (Dahl, 2004; Lupyan & Dale, 2010). In other words, we predict that having case marking will benefit comprehension despite the added burden of attending to it and producing it. As a secondary goal, we wanted to explore the impact of the redundant cue on producing other aspects of the language, in our case, word order: We wanted to see if learners of the redundant language will be more (or less) accurate in their production of the novel word order than learners of the control language. The predictions regarding word order accuracy in production are less clear-cut. On the one hand, if redundancy helps learners better master the entire linguistic system, it could lead to more accurate production of various aspects of the language. In our design, this would mean that learners of the redundant language will be more accurate at producing the correct word order than learners of the control language, and that using the redundant case marking cue will be associated with increased accuracy in producing the correct word order. Alternatively, the need to produce an additional element could make the sentences in the redundant language harder to produce. In our design this would mean that case marking would facilitate comprehension of thematic assignment, but it will not facilitate production of the correct word order. Such a dissociation between comprehension and production pressures is documented in other linguistic domains, such as semantic extension (Harmon & Kapatsinski, 2017) and regularization (Schwab, Lew-Williams, & Goldberg, 2018).

2. Method

2.1. Participants

60 children (age range: 7;0–9;0y, mean age: 7.10y, 19 females) and 56 adults (age range: 20–38y, mean age: 25y, 26 females) took part in the study. All child participants were visitors at the Bloomfield Science Museum in Jerusalem. They were recruited for this study as part of their visit to the Israeli Living Lab in exchange for a small reward. Parental

consent was obtained for all children. Adult participants were students at the Hebrew University (recruited for this study for credit or payment). All participants were native Hebrew speakers, and none of them had known language or learning disabilities. In each age group participants were randomly assigned to one of the two language conditions: the redundant language (30 child participants, 28 adult participants) and the control language (30 child participants, 28 adult participants).

2.2. Materials

In both conditions, participants were exposed to the same lexicon, which was composed of 6 semi-artificial Hebrew nouns (Hebrew nouns with nonce suffixes) and two Hebrew verbs. All nouns corresponded to masculine human characters, and indicated professions (e.g., clown, chef). The verbs were the Hebrew words of “kick” and “touch”. The constituent order of the language in both conditions was the non-Hebrew like OSV. In the redundant language a nonce case marking (“pazz”) followed all objects, while in the control language there was no such case marking. This cue was also non-native-like: Hebrew doesn't have post-nominal case-marking on objects. Sentences were concatenated from recordings of the individual words spoken by a female native Hebrew speaker. Participants saw and described the exact same drawings in both conditions.¹

2.3. Procedure

Participants were told they were going to meet some aliens who “say things differently from us” and that they would learn to speak like these aliens (adult participants were told that the experiment is designed for children, to explain the child-like instructions, but that the learning itself would not be easy). Participants sat with headphones in front of a computer. Child participants sat next to a research assistant, who provided them with verbal instructions. They saw drawings and heard recorded descriptions of these drawings in the alien-language. The experiment had several stages of training before reaching the crucial sentence production and comprehension stage (see Fig. 1). First, a noun exposure phase, in which participants saw each character, heard its name in the alien language, and had to repeat the name out loud (6 trials, one per noun). In both conditions, the noun label was presented without case marking. This was followed by a noun comprehension test (12 trials, two per noun) where participants saw two drawings, heard one label and had to match the label to the correct drawing. Feedback was provided after each trial. The following phase was sentence exposure (12 trials) where participants saw a drawing of a transitive action (involving two of the characters, all characters could appear as agents and patients), heard a transitive sentence, and had to repeat it. The position of the agent and the patient in the drawing (left vs. right) was counterbalanced.

The next stages assessed what participants had learned. The first was a sentence comprehension test (12 trials) where participants saw two drawings of events, heard a sentence, and had to match the sentence to the correct drawing. All the sentences here involved previously unheard combinations of agents and patients. The participants had to use the mouse to choose the matching drawing. No feedback was given. The next phase was a sentence production test (12 trials) where participants saw previously unseen drawings of transitive actions and had to describe them in the alien-language (that is, orally produce entire sentences themselves). Participants' descriptions were recorded. These two tests allowed us to estimate how well participants learned how to produce and comprehend the novel transitive constructions. Participants in the redundant language condition had an additional sentence forced-choice

¹ The drawings were based on the stimuli from Branigan, Pickering, and Cleland (2000), and were produced by Sara Rolando. They are courtesy of Kenny Smith and Jennifer Culbertson from the University of Edinburgh.

phase (12 trials) to ensure they noticed the case marking cue. In this phase, they saw a previously unseen drawing, heard two descriptions of it, and had to choose the correct one. One option had case marking (like the sentences they heard before) and one was without case marking (as in the control language). Participants had to say which sentence better described the drawing by pressing “1” or “2”, corresponding to the order in which the options were presented.

3. Results

3.1. Comprehension

To make sure the lexicon was learned well in both conditions, we first examined participants' performance in the noun comprehension test. Indeed, participants in both age groups learned the lexicon equally well (this was confirmed by Wilcoxon rank sum tests. Adults: control 99.7%, redundant 99.4%, $W = 406$, $p = 0.57$; children: control 96.6%, redundant 98.6%, $W = 419.5$, $p = 0.45$). We next turned to our variable of interest — participants' performance in the sentence comprehension test, to see if comprehension was improved in the redundant language. Fig. 2 shows the comprehension scores of each age group in each language condition. We used a mixed-effect logistic regression model to examine the effect of language condition and age group on sentence comprehension (using the glmer function in R software, D. Bates, Mächler, Bolker, & Walker, 2015, and the maximum random effect structure justified by the data that converged, Barr, Levy, Scheepers, & Tily, 2013). The dependent variable was accuracy on each trial. The model included fixed effects for condition (redundant vs. control, sum coded), age group (children vs. adults, sum coded), their interaction, and trial number (centered). The model included random intercepts for participants (see Table 1 for full model).

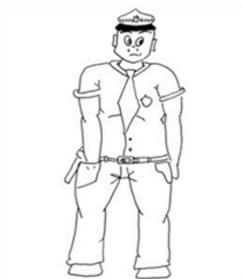
As predicted, comprehension scores were higher in the redundant language condition, indicating the redundant morpheme facilitated learning of thematic assignment (92% vs. 75%, $\beta = 0.84$, $SE = 0.21$, $p < 0.001$). In addition, adults received higher comprehension scores than children (90% vs. 78%, $\beta = 0.7$, $SE = 0.21$, $p = 0.001$), as is often the case in lab-based artificial language learning studies (e.g., Culbertson, Jarvinen, Haggarty, & Smith, 2019; Raviv & Arnon, 2018). The interaction between condition and age group was not significant ($\beta = -0.28$, $SE = 0.2$, $p = 0.17$). As can be seen in Fig. 2, however, child and adult learners do seem to differ in their comprehension patterns. While the difference between the two conditions seems small in the adult group, it is much larger in the children's group. To further investigate this, we conducted post-hoc simple effects models for each of the age groups separately. Each model included fixed effects for condition (redundant vs. control, sum coded), and trial number (mean-centered), and random intercepts for participants. The model for the children's group had an additional fixed effect for child's age (mean-centered, see Tables 2 and 3 for full models). While there was no difference between the conditions for adults (93% vs. 86%, $\beta = 0.74$, $SE = 0.48$, $p = 0.12$), children showed significantly better learning in the redundant language condition (91% vs. 65%, $\beta = 1.04$, $SE = 0.2$, $p < 0.001$). Importantly, the lack of difference for the adults should be interpreted with caution: Adults show high performance in both conditions, and are close to ceiling in the redundant condition, making it hard to tell whether the redundant morpheme could facilitate their learning (like the children).

3.2. Production

Participant's productions were transcribed and coded for word order accuracy and use of the case marker to see a) whether case marking was used in productions of the participants in the redundant condition, and b) whether having case marking in the language facilitated the production of the correct word order (i.e., whether word order was more accurately produced in the redundant language than in the control language). Both procedures were done by research assistants blind to the

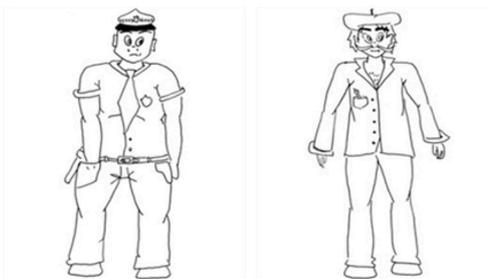
(A) Noun exposure

🔊 shoterig



(B) Noun comprehension test

🔊 shoterig



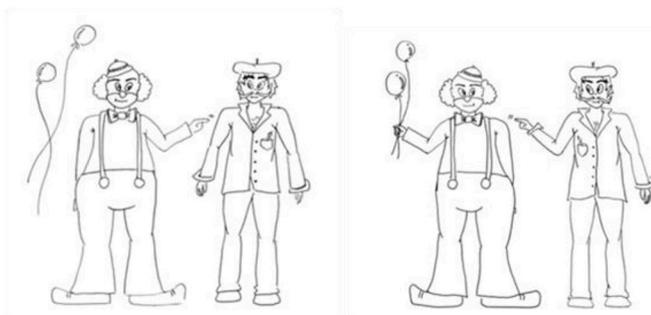
(C) Sentence exposure

🔊 Tabaxig(pazz) ganavig ba'at

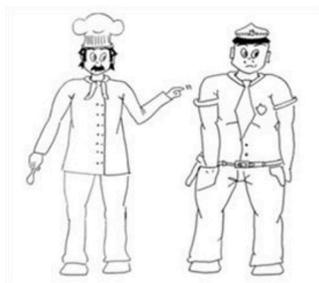


(D) Sentence comprehension test

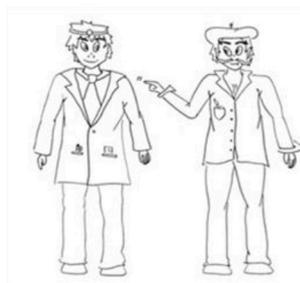
🔊 Zayarig(pazz) leizanig naga



(E) Sentence production test



(F) Sentence forced choice *



1 Rofeig zayarig naga

2 Rofeigpazz zayarig naga

Fig. 1. Experimental procedure (A–F). The sentence forced-choice phase (F) appeared only in the redundant language condition. The language was only aurally presented in the experiment. The written form appears here only for demonstration.

condition and the experimental hypothesis. Four participants in the adult group had no recordings due to experimental error, and one adult participant had recordings that were cut in the middle. Those 5 participants were excluded from the production analysis, leaving us with 25 adult participants in the control condition, and 26 in the redundant condition.

We begin with investigating whether participants in the redundant language condition learned the redundant cue itself. In line with our predictions, children and adults in the redundant language condition produced case marking in the majority of their productions (children: $M = 85\%$, $SD = 33\%$, adults: $M = 88\%$, $SD = 32\%$), indicating that they noticed and learned the cue, and treated it as an inherent part of the system, despite its redundancy and the additional effort involved in producing it. This pattern was further confirmed by looking at participants' responses in the sentence forced-choice phase (recall that only participants in the redundant language condition had this additional part). Both children and adults preferred sentences with case marking (this was confirmed by Wilcoxon rank sum tests. Children: $M = 88\%$, $SD = 20\%$, significantly greater than chance: $W = 421$, $p < 0.001$; adults: $M = 99\%$, $SD = 4\%$, significantly greater than chance: $W = 406$, $p <$

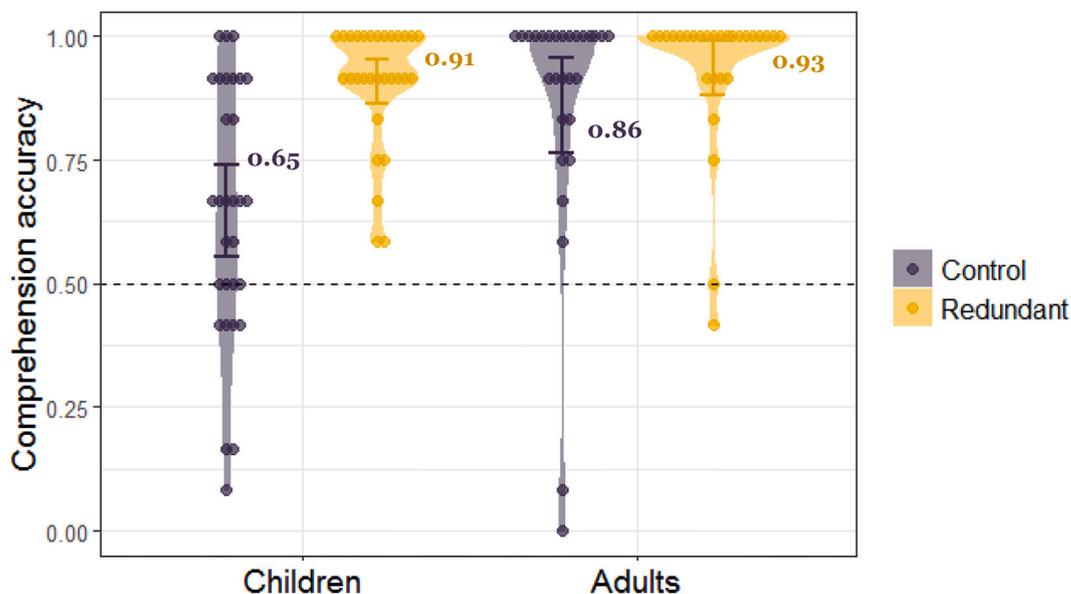


Fig. 2. Accuracy scores by language condition and age group. The dashed line indicates the chance level; error bars indicate confidence intervals of 95%; individual points indicate by-participant means; numbers represent means.

Table 1
Regression model for comprehension scores in both age groups.

	Estimate	Std. Error	z-value	p-value
Intercept	2.687	0.237	11.329	<0.001***
Condition = redundant	0.838	0.207	4.054	<0.001***
Group = adults	0.696	0.209	3.336	0.001***
Trial number	0.016	0.026	0.605	0.545
Condition = redundant * Group = adults	-0.278	0.203	-1.367	0.172

Note. Here and in all following tables, all fixed effects were sum-coded. The model's intercept therefore refers to the grand mean of all predictors. In this table, it reflects the log-odds of choosing the correct drawing when the predictor is the mean value of all fixed effects. The significant effect of the intercept indicates that mean performance was significantly above chance (50%).

Table 2
Regression model for comprehension scores in the adult group.

	Estimate	Std. Error	z-value	p-value
Intercept	4.344	0.741	5.863	<0.001***
Condition = redundant	0.736	0.478	1.54	0.123
Trial number	-0.03	0.049	-0.607	0.544

Table 3
Regression model for comprehension scores in the children group.

	Estimate	Std. Error	z-value	p-value
Intercept	1.801	0.212	8.501	<0.001***
Condition = redundant	1.036	0.204	5.082	<0.001***
Trial number	0.033	0.03	1.091	0.275
Age	-0.368	0.327	-1.123	0.261

0.001).

We next turned to compare the production of word order in the two conditions. Fig. 3 shows the proportion of producing the correct OSV word order for each age group in each language condition. We used a mixed-effect logistic regression model to examine the effect of language condition and age group on OSV production. The dependent variable was word-order accuracy on each trial (OSV vs. not OSV). The model

included fixed effects for condition (redundant vs. control, sum coded), age group (children vs. adults, sum coded), their interaction, and trial number as a centered continuous factor. Additionally, it included random intercepts for participants (see Table 4 for full model). We found no significant difference in word order accuracy between the two conditions, although the trend was in favor of the redundant language condition (88% vs. 76%, $\beta = 0.75$, $SE = 0.61$, $p = 0.22$). There was also no difference between the age groups (90% vs. 76%, $\beta = 0.93$, $SE = 0.61$, $p = 0.13$), and no interaction between condition and age group ($\beta = 0.12$, $SE = 0.61$, $p = 0.84$). Taken together, production of the correct word order did not seem to be facilitated in the redundant language condition for any of the age groups, but it did not seem to be impeded by it either.²

Finally, to gain more insight into the relation between the production of case marking and the correct word order, we looked at the production patterns of participants in the redundant language condition (since only they could potentially use both cues in production). As mentioned above, children and adults produced case marking in the majority of their productions. Interestingly, participants also mostly produced the correct word order (children: $M = 82\%$, $SD = 35\%$, adults: $M = 95\%$, $SD = 20\%$), indicating they managed to learn the non-Hebrew order. Fig. 4 shows the proportion of correct word order (OSV) and case marking

² To make sure the effect of condition is not obscured in our analyses because of the bimodality present in the data (see Fig. 3), we also tested the difference between the means using non-parametric permutation analysis (the medians were 1 in both age groups and conditions, rendering means as a better summary statistics). We conducted this analysis separately for each age group. This analysis consisted of several steps. First, we calculated the difference between the mean of learners' mean OSV use in each condition (i.e., the difference between mean OSV use for each age group in the control condition and the redundant condition as seen in Fig. 3). Next, we randomly assigned each participant's mean OSV use to one of the two conditions (i.e., we shuffled the link between production scores and condition). Following this shuffling, we recomputed the difference between the conditions' means. We repeated this process — shuffling and recomputing the difference between conditions' means — 10,000 times. Finally, we calculated the proportion of times where the difference between the means was larger in the permutations compared to our original results. This proportion amounts to the p-value. In both age groups, we found no significant difference in word order accuracy (mean OSV use) between conditions (children: $p = 0.23$, adults: $p = 0.15$).

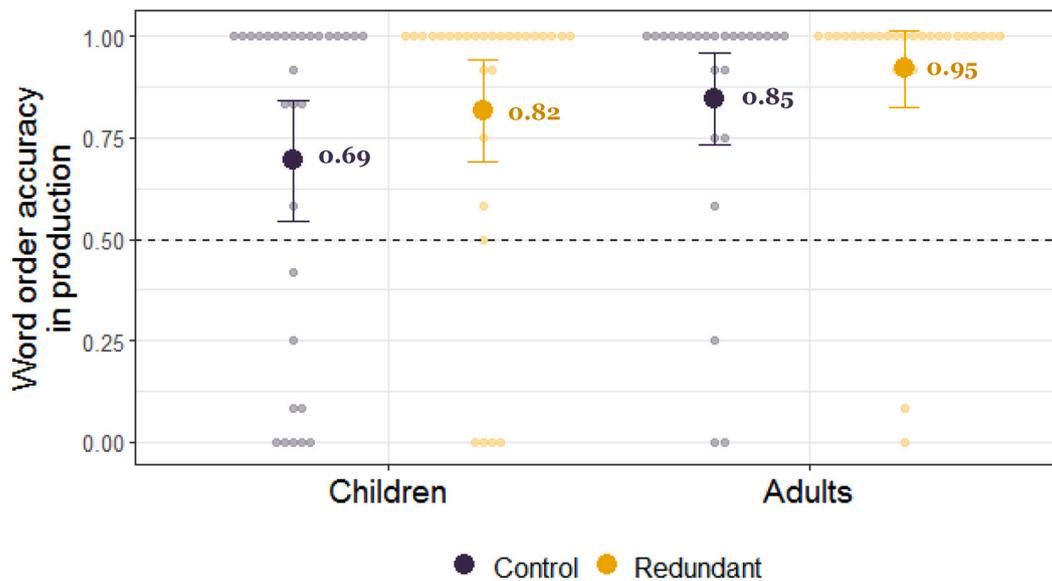


Fig. 3. Word order accuracy in production by language condition and age group. The dashed line indicates the chance level; error bars indicate confidence intervals of 95%; individual points indicate by-participant means; large points and numbers represent means.

Table 4
Regression model for OSV production scores in both age groups.

	Estimate	Std. Error	z -value	p-value
Intercept	3.484	0.612	5.69	<0.001 ***
Condition = redundant	0.749	0.612	1.224	0.221
Group = adults	0.926	0.612	1.512	0.13
Trial number	0.012	0.041	0.3	0.764
Condition = redundant * Group = adults	0.123	0.612	0.201	0.841

productions by participants in the redundant language condition. As can be seen in the plot, both children and adults were very consistent in their productions, with most learners always producing both case marking and the correct word order. This pattern could suggest a facilitative relation between the production of case marking and the correct word order. However, our data contains too few errors to test the possibility (or direction) of this possible relation: Learners, especially adult ones, simply did not make enough production errors to carry out a meaningful analysis. In addition, the plot illustrates that error rates were driven by very few participants: Most learners (children and adults) always produced both case marking and the correct word order (19/26 adults and 21/30 children). We return to this in the Discussion.

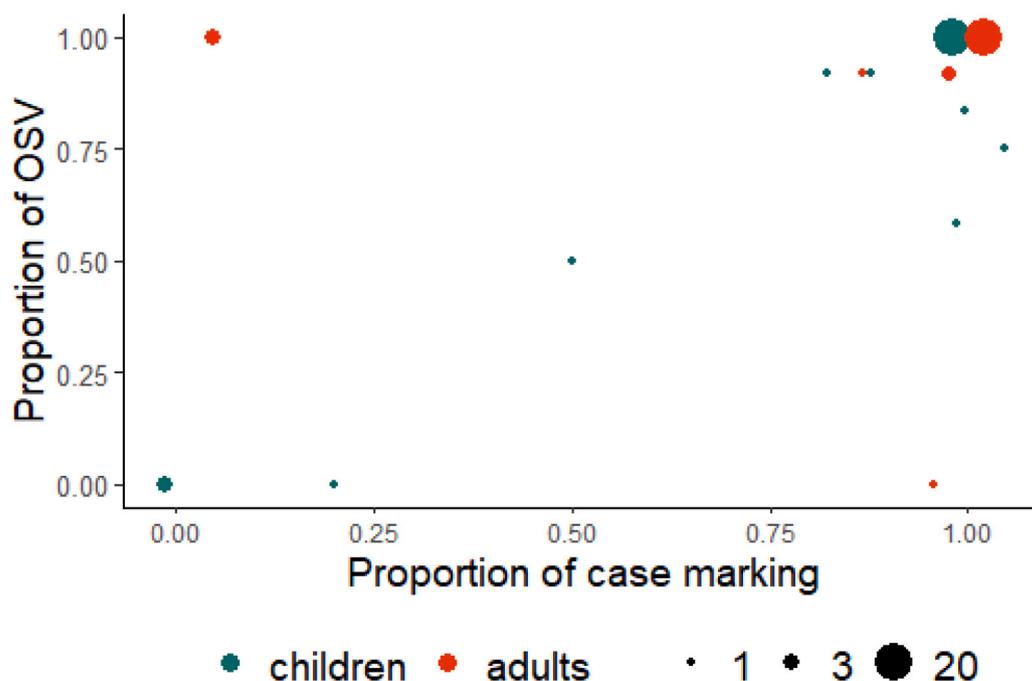


Fig. 4. Proportion of the correct word order (OSV) and case marking in productions of learners in the redundant language condition. Dot size represents number of participants.

4. Discussion

The presence of redundancy in languages calls for an explanation: If speakers are driven by a bias for efficient communication (Aylett & Turk, 2004; Cohen Priva, 2015; Jaeger, 2010; Kurumada & Jaeger, 2015; Levy & Jaeger, 2007; Pate & Goldwater, 2015; Piantadosi, Tily, & Gibson, 2011), then why do languages use more than one cue to convey the same information? Our study set out to test the prediction that redundant marking could be functional for learning, and may aid mastery of novel morpho-syntactic structures. In particular, we predicted, and found, that redundant marking can facilitate comprehension of thematic assignment despite the added burden of having to attend to, and produce, an additional cue. Children and adults learned two versions of an otherwise identical artificial language with fixed OSV word order: One version had a redundant morpheme to signal thematic assignment (the redundant language) and one did not (the control language). The redundant language was more complex than the control language, since it contained an additional cue to attend to and learn. Despite this, children learned the cue – as evidenced in them using it in their own productions and in their preferences in a forced-choice task – and were more accurate in comprehending the novel thematic assignment in the redundant language. Children showed similar accuracy in producing the correct word order in both conditions: The additional effort of producing the case marking cue did not impede (or facilitate) production of the correct word order. The lack of difference could be interpreted in different ways. It could indicate that redundancy doesn't facilitate production in the same way as it facilitates comprehension, because the additional cue makes sentences more challenging to produce overall. Alternatively, one could argue that even though sentences are harder to produce in the redundant language, word order accuracy was not impeded, pointing to an overall facilitative effect (comprehension is aided and production is not hurt). Additional work is needed to see whether redundant cues harm production accuracy of other linguistic aspects of a novel language (we further discuss the different pressures involved in production and comprehension below).

Interestingly, the facilitative effect of the redundant morpheme on comprehension was not detected for adult learners, who showed accurate comprehension of thematic assignment, as well as accurate production of word order in both language versions. It is hard to tell whether the lack of effect reflects adults' greater ease in learning the languages (they were close to ceiling at both measures) or a more fundamental difference in the impact of redundant cues on children and adults, as has been suggested by the LNH (Dale & Lupyan, 2012; Lupyan & Dale, 2010). In the current study, to make the two groups of learners as comparable as possible, children and adults completed the exact same task. This, however, meant that the task was easier for adults. Creating a more complex version for adults could help us clarify whether they indeed benefit less from redundant cues, but would prevent the direct comparison between them and child learners. While various researchers agree that adult learning is a key factor in explaining the simplification of language (Bentz & Winter, 2013; Dahl, 2004; Nettle, 2012; Trudgill, 2011; Wray & Grace, 2007), the question of what gives rise to complexity in the first place is less clear (Lupyan & Dale, 2010; Trudgill, 2011). Under the LNH account, the answer lies in child learners: Child learners are predicted to benefit from redundancy, leading in turn to the creation and preservation of redundancy for languages used by communities with more child learners. Importantly, the current findings support the novel prediction made by the LNH. Future work is needed to further compare the impact of redundancy on learning between children and adults. One way to go about this is to create a more complex version that is still learnable by children. Another, is to use a design where learning rates can be assessed and compared between the age groups (i. e., children may show faster learning rates in the redundant condition than in the control while adults may not). Finding that children benefit from redundant cues more than adults, even in a language that is harder to learn, would have important consequences for the LNH and for

theories of language change more generally. Such a pattern would join others pointing to differences in the learning biases of children and adults (Culbertson et al., 2019; Gopnik, 2020; Hudson Kam & Newport, 2005, 2009), and the impact of these differences on language emergence and change (Labov, 2007; Raviv & Arnon, 2018).

The current study raises additional questions to explore in future work. One is whether the same facilitative effects can be found in other linguistic domains beyond thematic assignment. The phenomenon of multiple exponence, where a single meaning is redundantly marked using multiple morphological markers, is an example in point. A recent study has found that multiple exponence can facilitate language processing, but only when listeners expect it to occur (Caballero & Kapatsinski, 2015). This study was conducted on speakers of Choguita Rarámuri, a language where some meanings are cued by two consecutive suffixes, the second of which is optional. Participants had to recognize words that either had or did not have a redundant suffix, embedded in different levels of acoustic noise. In addition, the experimental stimuli were constructed such that some words were generally easier to recognize in context than others. They found that the redundant morpheme aided recognition when words were generally more difficult to recognize, but made recognition worse when words were generally easier to recognize. The authors concluded that redundancy can benefit processing as a function of speakers' expectations, which are driven in turn by communicative efficient language use (Caballero & Kapatsinski, 2015). That is, redundancy is expected to appear only when understandability is otherwise compromised, and therefore speakers expect to find redundant forms only when they are needed. However, if redundancy benefits learning, then just like in the current study, multiple exponence might aid learning of *novel* structures. In ongoing work, we investigate this possibility using an artificial language learning paradigm.

An additional open question has to do with the prediction that redundancy is found more when conversing with learners. In an additional line of work we show that interactions with learners are indeed characterized by more redundancy (Tal, Grossman, & Arnon, 2021; Tal, Grossman, Rohde, & Arnon, 2021). In these studies, we investigate redundancy at the discourse and full-language level. If listeners' language proficiency impacts the redundancy in speakers' speech, this should be evident also in morpho-syntactic redundancy. Future work could investigate, for example, whether speakers of languages with optional case marking use more case marking when talking to language learners.

Finally, the current study shows that learning from two linguistic cues can facilitate comprehension. However, we do not know the exact cue structure that was learned. That is, we do not know whether learners learned both cues as two converging cues, or as one "rich" cue. This question is further motivated by our finding that most learners in the redundant language condition consistently used both case marking and the correct word order in their productions. We investigate this question in other work by dissociating between the two cues in exposure and testing (Tal & Arnon, in prep). In line with proposals in the language acquisition literature (Ibbotson & Tomasello, 2009), we find that while benefiting learning of thematic assignment, when children learn both cues at once, it hinders their ability to identify the contribution of each cue alone. This highlights an important feature of the language learned in the current study: The redundant cues were *deterministic*, they were always present (perfectly available, using the terminology of the Competition Model). Importantly, however, redundant cues in natural languages are often probabilistic (E. Bates & MacWhinney, 1989; Levshina, 2020; Monaghan, 2017; Monaghan, Brand, & Frost, 2017; Monaghan & Christiansen, 2008). In fact, it has been argued that part of the reason that languages have redundant cues is to compensate for the fact that individual cues are often not deterministic (Monaghan, 2017; Monaghan et al., 2017). In line with this reasoning, when redundant cues appear in learning but not in testing, learning is better facilitated when the cues are probabilistic rather than deterministic (Monaghan,

2017; Monaghan et al., 2017). These results make intuitive sense: If a learner learns a perfectly available cue (one that occurs in each case it is “needed”), but then this cue suddenly disappears, this could impair learning, since the cue may be relied upon as part of the structure learned. When cues are probabilistic, on the other hand, the learned structure is not exclusively associated with the cues, therefore allowing for their absence (Monaghan, 2017). These findings point to an intricate interplay between the advantage multiple cues hold for learning and the over-reliance they might create when they are too strong.

More broadly, the current findings have implications for our understanding of how language structure in general, and redundancy in particular, are shaped by both communicative and learning biases. Although in many languages the input children hear contains multiple cues for thematic assignment (Dittmar et al., 2008; Ibbotson & Tomasello, 2009), typological studies suggest that languages tend to trade-off these two cues: Languages that rely on fixed word order to encode thematic assignment often lack productive case marking (Koplenig, Meyer, Wolfer, & Mu, 2017; Siewierska, 1998; though see Levshina, 2020). This process can be documented historically in some languages (e.g., Old English, Marchand, 1951, though see Detges, 2009; Pintzuk, 2002 for challenges to these claims). Recent experimental work suggests this trade-off reflects speakers' bias for efficient communication (Fedzechkina & Jaeger, 2020; Fedzechkina, Newport, & Jaeger, 2017; Roberts & Fedzechkina, 2018). Using artificial language learning paradigms, Fedzechkina and colleagues found that learners preserved case marking in their language when word order was flexible (and there was more ambiguity), but decreased case marking when it was fixed. They suggest that learners aim to *avoid redundancy* by using case marking only when it is necessary. On the surface, these findings seem to contrast with our own: In our study, learners kept (and used) a redundant morpheme even though it was not necessary for comprehension. However, this highlights the differential impact various pressures can have on language. First, while case marking was deterministic in our design, it was optional (appearing on 67% of objects) in Fedzechkina et al. (2017). The different production patterns found in both experiments are congruent with previous findings on the impact input structure has on production of case marking (K. Smith & Culbertson, 2020). Specifically, it is quite likely that participants in our design were trying to faithfully reproduce the patterns in their input (see also discussion in Fedzechkina et al., 2017). Another important factor to consider is the difference between pressures impacting production and comprehension (e.g., Harmon & Kapatsinski, 2017). While redundant morphemes can facilitate comprehension, they are costly during production³ (Fedzechkina & Jaeger, 2020; Kanwal, Smith, Culbertson, & Kirby, 2017; Kurumada & Jaeger, 2015; Zipf, 1949). In other words, even when redundancy benefits learning, it can still be dispreferred in production. For this reason, redundant morphemes might be more prevalent in speech directed to learners, but less so otherwise. An additional factor which can interact with these pressures, which the current results might be suggestive of, is the possible difference between child and adult learners. Specifically with respect to cues to thematic assignment, redundant cues seem to be prevalent in child-directed speech (Dittmar et al., 2008; Ibbotson & Tomasello, 2009). If adult learners do not benefit from this redundancy to the same extent (Dale & Lupyan, 2012; Lupyan & Dale, 2010), then speech directed to adult learners might have less redundant cues for

³ There is anecdotal evidence in Fedzechkina et al. (2017) suggesting that although participants decreased the use of redundant case marking in their productions, the redundant case marking was nonetheless beneficial for learning: Prior to participants exclusion, thematic assignment was learned better in the redundant language version compared to the non-redundant version. Importantly, however, in this experiment the different versions rely on difference in word order, making the redundant language more similar to English - participants' native language. A more systematic investigation should therefore be in place.

thematic assignment. This is in line with the finding that languages with more adult learners tend to lose case marking (Bentz & Winter, 2013). In other words, the observed trade-off between case marking and word order seen in many languages might reflect languages' adaptation to learning biases of adult learners (Bentz & Winter, 2013; Lupyan & Dale, 2010; Trudgill, 2011; Wray & Grace, 2007). Future experimental work is needed to investigate the differences between child and adult learners and the impact these differences might have on redundant morphology.

In sum, the current findings suggest that redundancy of linguistic cues can be facilitative in certain learning situations. This is in line with previous work on the effect of learning from multiple cues (Monaghan, 2017; Sloutsky & Robinson, 2013; Smith, Colunga, & Yoshida, 2010). The present study serves as an important step toward understanding the functionality of redundancy in languages. Our findings show redundant cues can help comprehension of novel constructions even when these cues need to be learned and reproduced as part of the system.

CRediT authorship contribution statement

Shira Tal: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – original draft. **Inbal Arnon:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Acknowledgments

We wish to thank Henry Brice, Ram Frost, Noam Siegelman and Kenny Smith for helpful discussions. We also want to thank Rimon Alyagon Dar and Tali Cohen for their help in collecting the data. We thank the Living Lab staff and the Bloomfield Science Museum in Jerusalem, as well as the parents and children who participated. The research was funded by the Israeli Science Foundation grant number 584/16 awarded to the second author. The first author was funded by the Mandel Scholion Center at the Hebrew University of Jerusalem.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2022.105055>.

References

- Audring, J. (2014). Gender as a complex feature. *Language Sciences*, 43, 5–17. <https://doi.org/10.1016/j.langsci.2013.10.003>
- Aylett, M., & Turk, A. (2004). The Smooth Signal Hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47(1), 31–56.
- Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, 36(2), 190–201. <https://doi.org/10.1037/0012-1649.36.2.190>
- Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science*, 13(3), 99–102.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney, & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–76). Cambridge University Press.
- Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, 11(3), 245–299. [https://doi.org/10.1016/0010-0277\(82\)90017-8](https://doi.org/10.1016/0010-0277(82)90017-8)
- Bentz, C., Ruzsics, T., Koplenig, A., & Samardzic, T. (2016). A comparison between morphological complexity measures: Typological data vs. language corpora. In *Proceedings of the workshop on computational linguistics for linguistic complexity* (pp. 142–153). <https://www.aclweb.org/anthology/papers/W/W16/W16-4117/>.
- Bentz, C., Verkerk, A., Kiela, D., Hill, F., & Buttery, P. (2015). Adaptive communication: Languages with more non-native speakers tend to have fewer word forms. *PLoS One*, 10(6), 1–23. <https://doi.org/10.1371/journal.pone.0128254>
- Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change*, 3, 1–27. <https://doi.org/10.1163/22105832-13030105>

- Brnanigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75, 13–25.
- 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 and Language*, 89, 68–86. <https://doi.org/10.1016/j.jml.2015.12.009>
- Caballero, G., & Harris, A. C. (2012). A working typology of multiple exponence. In F. Kiefer, M. Ladányi, & P. Siptár (Eds.), *Current issues in morphological theory: (Ir) regularity, analogy and frequency. Selected papers from the 14th International Morphology Meeting, Budapest* (pp. 163–188). John Benjamins Publishing.
- Caballero, G., & Kapatsinski, V. (2015). Perceptual functionality of morphological redundancy in Choguita Rarámuri (Tarahumara). *Language, Cognition and Neuroscience*, 30(9), 1134–1143. <https://doi.org/10.1080/23273798.2014.940983>
- Chan, A., Lieven, E., & Tomasello, M. (2009). Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics*, 20(2), 267–300. <https://doi.org/10.1515/COGL.2009.015>
- Christiansen, M. H., & Monaghan, P. (2006). Discovering verbs through multiple-cue integration. In K. Hirsh-Pasek, & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 88–107). Oxford University Press.
- Cohen Priva, U. (2015). Informativity affects consonant duration and deletion rates. *Laboratory Phonology*, 6(2), 243–278. <https://doi.org/10.1515/lp-2015-0008>
- Culbertson, J., Jarvinen, H., Haggarty, F., & Smith, K. (2019). Children's sensitivity to phonological and semantic cues during noun class learning: Evidence for a phonological bias. *Language*, 95(2), 268–293. <https://doi.org/10.1353/lan.2019.0031>
- Dahl, Ö. (2004). *The growth and maintenance of linguistic complexity*. John Benjamins Publishing Company.
- Dale, R., & Luppian, G. (2012). Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems*, 15(3), 1150017–1–1150017–16.
- Detges, U. (2009). How useful is case morphology? The loss of the old French two-case system within a theory of preferred argument structure. In J. Barðdal, & S. Chelliah (Eds.), *The role of semantic, pragmatic, and discourse factors in the development of case* (pp. 93–120). John Benjamins Publishing.
- Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development*, 79(4), 1152–1167.
- 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.
- Fedzechkina, M., Newport, E. L., & Jaeger, T. F. (2017). 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>
- Flom, R., & Bahrick, L. E. (2010). The effects of intersensory redundancy on attention and memory: Infants' long-term memory for orientation in audiovisual events. *Developmental Psychology*, 46(2), 428–436. <https://doi.org/10.1037/a0018410>
- Frank, A. F., & Jaeger, T. F. (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. In *The 30th annual meeting of the cognitive science society* (pp. 939–944).
- Gerken, L. A., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32(2), 249–268. <https://doi.org/10.1017/S0305000904006786>
- Gibson, E., Futrell, R., Piantadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, 1–19. <https://doi.org/10.1016/j.tics.2019.02.003>
- Givón, T. (1991). Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language*, 15(2), 335–370. <https://doi.org/10.1075/sl.15.2.05giv>
- Gopnik, A. (2020). Childhood as a solution to explore – Exploit tensions. *Philosophical Transactions of the Royal Society B*, 375(1803), 20190502.
- Haig, G., & Forker, D. (2018). Agreement in grammar and discourse: A research overview. *Linguistic Typology*, 56(4), 715–734.
- Harmon, Z., & Kapatsinski, V. (2017). Putting old tools to novel uses: The role of form accessibility in semantic extension. *Cognitive Psychology*, 98, 22–44. <https://doi.org/10.1016/j.cogpsych.2017.08.002>
- Harris, A. (2017). *Multiple exponence*. Oxford University Press.
- Haspelmath, M. (2021). Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics*, 1–29 (Lingbuzz/004531%0A).
- Hengeveld, K., & Leufkens, S. (2018). Transparent and non-transparent languages. *Folia Linguistica*, 52(1), 139–175.
- Holler, J., & Levinson, S. C. (2019). Multimodal language processing in human communication. *Trends in Cognitive Sciences*, 23(8), 639–652. <https://doi.org/10.1016/j.tics.2019.05.006>
- Hudson Kam, C. L., & 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/10.1207/s15473341l1d0102_3
- Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, 59(1), 30–66. <https://doi.org/10.1016/j.cogpsych.2009.01.001>
- Ibbotson, P., & Tomasello, M. (2009). Prototype constructions in early language acquisition. *Language and Cognition*, 1, 59–85. <https://doi.org/10.1515/LANGCOG.2009.004>
- Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61(1), 23–62. <https://doi.org/10.1016/j.cogpsych.2010.02.002>
- Jaeger, T. F., & Buz, E. (2017). Signal reduction and linguistic encoding. In E. Fernández, & H. Cairns (Eds.), *The handbook of psycholinguistics* (pp. 38–81). John Wiley & Sons.
- 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.
- Koplenig, A., Meyer, P., Wolfer, S., & Mu, 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). <https://doi.org/10.7910/DVN/8KH0GB>
- Kravtchenko, E. (2014). Predictability and syntactic production: Evidence from subject omission in Russian. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 785–790). http://people.ucsc.edu/~ekravtch/Kravtchenko_CogSci2014.pdf
- 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>
- Labov, W. (2007). Transmission and diffusion. *Language*, 83(2), 344–387. <https://doi.org/10.1353/lan.2007.0082>
- Leufkens, S. (2020). A functionalist typology of redundancy. *Revista Da Abralin*, 19(3), 79–103.
- Levshina, N. (2020). Efficient trade-offs as explanations in functional linguistics: Some problems and an alternative proposal. *Revista Da Abralin*, 19(3), 50–78.
- Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In B. Schilöckopf, J. Platt, & T. Hoffman (Eds.), *vol. 19. Advances in neural information processing systems (NIPS)* (pp. 849–856). MIT Press.
- Lew-Williams, C., Ferguson, B., Abu-Zhaya, R., & Seidl, A. (2019). Social touch interacts with infants' learning of auditory patterns. *Developmental Cognitive Neuroscience*, 35, 66–74. <https://doi.org/10.1016/j.dcn.2017.09.006>
- Luppian, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS One*, 5(1). <https://doi.org/10.1371/journal.pone.0008559>
- MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308).
- MacWhinney, B. (2001). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 69–90). Cambridge University Press.
- Mahowald, K., Dautriche, I., Braginsky, M., & Gibson, E. (2020). *Efficient communication and the organization of the lexicon*.
- Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, 126(2), 313–318. <https://doi.org/10.1016/j.cognition.2012.09.010>
- Marchand, H. (1951). The syntactical change from inflectional to word order system and some effects of this change on the relation verb-object in English. *Anglia*, 70, 70–89.
- Matsuo, A., Kita, S., Shinya, Y., Wood, G. C., & Naigles, L. (2012). Japanese two-year-olds use morphosyntax to learn novel verb meanings. *Journal of Child Language*, 39(3), 637–663. <https://doi.org/10.1017/S0305000911000213>
- Monaghan, P. (2017). Canalization of language structure from environmental constraints: A computational model of word learning from multiple cues. *Topics in Cognitive Science*, 9, 21–34. <https://doi.org/10.1111/tops.12239>
- Monaghan, P., Brand, J., & Frost, R. L. A. (2017). Multiple variable cues in the environment promote accurate and robust word learning. In G. Gunzelman, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society (CogSci 2017)* (pp. 817–822).
- Monaghan, P., & Christiansen, M. H. (2008). Integration of multiple probabilistic cues in syntax acquisition. In *Corpora in language acquisition research: History, methods, perspectives* (pp. 139–164). John Benjamins.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, 19(4), 498–550. [https://doi.org/10.1016/0010-0285\(87\)90017-X](https://doi.org/10.1016/0010-0285(87)90017-X)
- Murgiano, M., Motamedi, Y., & Vigliocco, G. (2020). Language is far less arbitrary than one thinks: Iconicity and indexicality in real-world learning and processing. *Journal of Cognition*. <https://doi.org/10.5334/joc.113>
- Nettle, D. (2012). Social scale and structural complexity in human languages. *Philosophical Transactions of the Royal Society*, 367, 1829–1836. <https://doi.org/10.1098/rstb.2011.0216>
- O'Shanessy, C. (2010). Competition between word order and case-marking in interpreting grammatical relations: A case study in multilingual acquisition. *Journal of Child Language*, 38(2011), 763–792. <https://doi.org/10.1017/S0305000910000358>
- Pate, J. K., & Goldwater, S. (2015). Talkers account for listener and channel characteristics to communicate efficiently. *Journal of Memory and Language*, 78, 1–17.
- Piantadosi, S. T., Tily, H., & Gibson, E. (2011). The communicative function of ambiguity in language. *Cognition*, 122(3), 280–291. <https://doi.org/10.1016/j.cognition.2011.10.004>
- Pintzuk, S. (2002). Morphological case and word order in Old English. *Language Sciences*, 24, 381–395.
- Raviv, L., & Arnon, I. (2018). Systematicity, but not compositionality: Examining the emergence of linguistic structure in children and adults using iterated learning. *Cognition*, 181(August), 160–173. <https://doi.org/10.1016/j.cognition.2018.08.011>
- Roberts, G., & Fedzechkina, M. (2018). Social biases modulate the loss of redundant forms in the cultural evolution of language. *Cognition*, 171, 194–201. <https://doi.org/10.1016/j.cognition.2017.11.005>

- Schwab, J. F., Lew-Williams, C., & Goldberg, A. E. (2018). When regularization gets it wrong: Children over-simplify language input only in production. *Journal of Child Language*, 1–19. <https://doi.org/10.1017/S0305000918000041>
- Siewierska, A. (1998). Variation in major constituent order: A global and a European perspective. In A. Siewierska (Ed.), *Constituent order in the languages of Europe* (pp. 475–552) (Mouton De Gruyter).
- Sloutsky, V. M., & Robinson, C. W. (2013). Redundancy matters: Flexible learning of multiple contingencies in infants. *Cognition*, 126(2), 156–164. <https://doi.org/10.1016/j.cognition.2012.09.016.Redundancy>
- Smith, K., & Culbertson, J. (2020). *Communicative pressures shape language during communication (not learning): Evidence from casemarking in artificial languages*. <https://doi.org/10.31234/osf.io/5nwhq>
- Smith, L. B., Colunga, E., & Yoshida, H. (2010). Knowledge as process: Contextually cued attention and early word learning. *Cognitive Science*, 34(7), 1287–1314. <https://doi.org/10.1111/j.1551-6709.2010.01130.x>
- Tal, S., Grossman, E., & Arnon, I. (2021). Infant-directed speech becomes less redundant as infants grow: implications for language learning. <https://psyarxiv.com/bgtzd/>.
- Tal, S., Grossman, E., Rohde, H., & Arnon, I. (2021). Speakers use more redundant referents with language learners: evidence for communicatively-efficient referential choice. <https://psyarxiv.com/cw2be>.
- Thiessen, E. D., & Saffran, J. R. (2009). How the melody facilitates the message and vice versa in infant learning and memory. *Annals of the New York Academy of Sciences*, 1169, 225–233. <https://doi.org/10.1111/j.1749-6632.2009.04547.x>
- Trudgill, P. (2009). Sociolinguistic typology and complexification. In P. Trudgill, G. Sampson, & D. Gil (Eds.), *Language complexity as an evolving variable* (pp. 98–109). Oxford University Press.
- Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press.
- Wit, E. C., & Gillette, M. (1999). What is linguistic redundancy?. Technical Report <http://www.math.rug.nl/~ernst/linguistics/redundancy3.pdf>.
- Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117, 543–578. <https://doi.org/10.1016/j.lingua.2005.05.005>
- Yoshida, H., & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science*, 16(2), 90–95. <https://doi.org/10.1111/j.0956-7976.2005.00787.x>
- Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley Press.