Studies in Second Language Acquisition (2025), **47**, 1250–1276
doi:10.1017/S027226312510106X

RESEARCH ARTICLE

Learning morphology from cross-situational statistics

Liuqi Zhu[1], Patrick Rebuschat[1,2], Jessie S. Nixon[3] and Padraic Monaghan[4]

[1]Department of Linguistics and English Language, Lancaster University, UK; [2]LEAD Graduate School and Research Network, University of Tübingen, Germany; [3]Department of Linguistics, University of Bielefeld, Germany and [4]Department of Psychology, Lancaster University, UK
Corresponding author: Liuqi Zhu; Email: l.zhu8@lancaster.ac.uk

(Received 18 October 2024; Revised 16 May 2025; Accepted 17 June 2025)

Abstract

Non-native languages tend to be acquired through a combination of explicit and implicit learning, where implicit learning requires coordination of language information with referents in the environment. In this study, we examined how learners use both language input and environmental cues to acquire vocabulary and morphology in a novel language and how their language background influences this process. We trained 105 adults with native languages (L1s) varying in morphological richness (English, German, Mandarin) on an artificial language comprising nouns and verbs with morphological features (number, tense, and subject-verb [SV] agreement) appearing alongside referential visual scenes. Participants were able to learn both word stems and morphological features from cross-situational statistical correspondences between language and the environment, without any instruction. German-speakers learned SV agreement worse than other morphological features, which were acquired equally effectively by English or Mandarin speakers, indicating the subtle and varied influence of L1 morphological richness on implicit non-native language learning.

Keywords: bottleneck hypothesis; cross-linguistic influence; cross-situational learning; morphology; statistical learning

Introduction

The acquisition of non-native languages tends to involve a combination of explicit instruction and implicit acquisition of the associations between sentences and referents in the environment around the speaker (Rebuschat, 2022; Williams & Rebuschat, 2023). In this study, we unpack the way in which learners combine information from the language with features of the environment to acquire both words and functional morphemes. Learning to map words onto co-occurring referents in the environment poses an impressive challenge. It is often difficult to figure out the meaning of a novel

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

word based on one scene due to the potential ambiguity of possible mappings that can be made (Quine, 1960). However, recent research has suggested a way to solve the problem of referential ambiguity. After being exposed to multiple scenes, learners can determine the mapping between the sound and its referent, by keeping track of cross-situational statistics between words and referents that regularly appear together (Schroer & Yu, 2023; Smith & Yu, 2008; Yu et al., 2021). To date, cross-situational learning (CSL) is evidenced to be effective for learning referents that can be directly observed in the environment, such as nouns (e.g., Ge, Monaghan & Rebuschat, 2025; Suanda & Namy, 2012; Vlach & DeBrock, 2019) and verbs (e.g., Monaghan et al., 2015; Scott & Fisher, 2012). However, the linguistic features that pose most challenges to adult language learners often cannot be easily identified from the visual world, such as functional morphology (Slabakova, 2014). The question of whether adult language learners can learn morphology via CSL has to date only been addressed by very limited empirical studies.

Learning morphemes via CSL

As one of the few studies that explore how morphological features may be learned via CSL, Finley (2023) tested the extent to which morphological cues relating to the semantic categories of nouns could be acquired through CSL. When suffixes corresponded to semantic categories (e.g., animals, fruits, vehicles), the morphological cue could be rapidly acquired from CSL; furthermore, the cue could then be used to improve the learning of the stems (see also Monaghan & Mattock, 2012). The beneficial effect of the morphological cues for learning, however, was only realized when referential ambiguity was initially low, enabling the role of the stem and the morpheme to be discerned. Finley's (2023) Experiment 1 demonstrated that morphological learning from presentations of single words is, in principle, a possibility. However, the complexity of natural language learning involves determining the role of word stems and affixes within longer and more complex sentences. Sentences in the natural language learning environment also have richer syntax or morphology, potentially without the initially low referential ambiguity that Finley (2023) showed was required to facilitate CSL of morphology.

Reflecting one aspect of the complexity of natural language learning in terms of multiple words and referents occurring simultaneously, Rebuschat et al. (2021) adopted a CSL paradigm where participants were exposed to presentations of sentences with visual scenes. Specifically, they investigated whether the marker words indicating the agent and patient of the sentences could be acquired from complex sentence-scene correspondences. In their Experiment 1, adult participants were presented with a single scene comprising two aliens interacting with one another in each learning trial, which co-occurred with hearing a transitive sentence in an artificial language that described the scene. After exposure, learning of nouns, verbs, adjectives, and marker words as well as sensitivity to word order was then tested. Rebuschat and colleagues (2021) found that case markers can be learned but it was the least and latest learned linguistic features compared to nouns, verbs, and adjectives. In a follow-up experiment (their Experiment 2), they found that the learning of the case markers was only at a chance level when referential ambiguity was increased by presenting two different scenes (one that matched the sentence they heard and one foil) in each learning trial. This finding was consistent with results from Walker et al. (2020) and Monaghan et al. (2021), which used the same artificial language paradigm, and found only low levels of case marker

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

learning. Taken together, there is evidence that learning the mappings of speech that expresses relations to more abstract, and harder to observe, properties of the environment to which the language relates, under certain circumstances, is possible but only when the referential ambiguity stays low. However, the question remains why these markers are more challenging to acquire than other aspects of the language (e.g., Monaghan et al., 2021; Rebuschat et al., 2021; Walker et al., 2020).

Explanations for why morphology poses learning difficulties

There are several explanations for why morphology is more difficult to learn, which could operate individually or cumulatively to explain learning (DeKeyser, 2005; Ellis, 2022; Slabakova, 2014), and that could limit the effectiveness of CSL to support the acquisition of morphology. First, the immediacy, or transparency, of a cue's referent may influence natural language learning (Hofweber et al., 2023; Sehyr & Emmorey, 2019). In Rebuschat et al.'s (2021) study, the function of the case markers was not immediately available within a single scene but instead had to be interpreted from the interoperation among words within the sentence and between potential agent and patient actors in the environment. This meant that the case markers were more opaque as a referent to the markers than were the nouns, verbs, or adjectives because of their lower visual salience (van Zoest & Donk, 2005). If opacity of referents and mappings are contributors to CSL of morphology, then this is unlikely to be to the same degree across all targets for morphological features. For instance, tense morphology may be difficult to acquire because of difficulties in isolating temporal order in events (Tünnermann & Scharlau, 2018), whereas number (e.g., singular or plural) may be easier to acquire because of its more immediate appearance in environmental stimuli—if there is one or more than one panda in the environment then that is easily observed, but whether the panda walked yesterday or will walk tomorrow cannot be determined from observation of the (current) environment. In contrast, subject-verb (SV) agreement may be hard to acquire both because it is not visually or contextually apparent in the environment and also because of the integration of syntactic and morphological constraints that are required, potentially explaining why SV agreement is widely observed to be difficult to learn in non-native language acquisition studies (Slabakova, 2014).

A second potential explanation for why morphology might be more difficult to acquire is that morphological cues are usually not acoustically salient in a speech stream (e.g., syllabically shorter with more reduced vowels than word stems) (Gass et al., 2018). Due to the limited capacity of working memory, learners tend to rely on the more salient cues (Cintrón-Valentín & Ellis, 2016). In Rebuschat et al. (2021), for instance, the case markers were shorter than the nouns and verbs (monosyllabic versus bisyllabic). Participants might have focused more on the learning of nouns and verbs as they are perceptually easier to process. However, the low level of learning of case markers in Rebuschat et al. (2021) could have been due to many possibilities, including cue salience, but also potentially the opacity or transparency of the morphological cues. Our study is designed to tease apart some of these contributors to difficulties and differences in learning morphology.

A third possible contributor to morphology learning difficulty is the influence of learners' native languages (L1s). Learning an additional language can be influenced by previously learned languages, a phenomenon referred to as cross-linguistic influence (Suethanapornkul, 2020) and may involve competition among cues available in known and additional languages (Nixon, 2020). A multitude of empirical evidence shows that

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

L1 plays a dominant role in influencing additional language learning, both in classroom (Choi & Ionin, 2021; Finn & Hudson Kam, 2015) and immersion (Diaubalick & Guijarro-Fuentes, 2019) settings. For instance, in an implicit language learning task, Ellis (2007) found that adult native speakers of an L1 that has little inflectional morphology tend to pay more attention to lexical cues than to morphological cues that encode the same meaning. The information that morphological features carry is thus, at least in some cases, redundant and this could result in the learning of the morphological cues being blocked by cues that are more learnable to the participants, which also varies based on L1 (see Nixon, 2020, for a discussion of blocking in speech learning). Similarly, a large dataset of non-native Dutch learners revealed that those adults with morphologically less complex L1s performed worse in acquiring Dutch than those with morphologically richer L1s (van der Slik, van Hout & Schepens., 2019). Hence, the speaker's L1 may influence the degree to which different morphological features are detected by the speaker.

Overall, cumulated evidence seems to suggest that there is a transfer from the richness of L1 morphology. However, the extent to which transfer occurs has been proposed to depend on structural linguistic similarity on a feature-by-feature basis, rather than wholesale effects from previously learned language (e.g., Slabakova, 2017; Westergaard et al., 2017). In particular, similarities of morphological features have been widely documented as affecting non-native language learning (Gardner et al., 2021; Hawkins & Hattori, 2006). For example, Mandarin indicates tense with adverbs or prepositions instead of inflectional cues (e.g., 他一般周一游泳 *tā yībān zhōuyī yóuyǒng*, "He usually Monday swim"; 我明天学习 *wǒ míngtiān xuéxí*, "I tomorrow study"), whereas English and German tend to indicate tense using morphology or short auxiliary verbs (e.g., *Ich gehe*, "I walk"; *Ich werde gehen*, "I will walk"), sometimes in addition to prepositions. English and German also indicate number (singular or plural) using a suffix morpheme. Mandarin, in contrast, only indicates number using one of several noun-specific classifiers preceding the noun (e.g., plum, Pflaume, 李子 "lǐzi"; plums, Pflaumen, 一袋李子 "Yī dài lǐzi" "a bag of plums"). When learning morphological cues for tense and number in a non-native language, if there is feature-by-feature transfer, L1 English and German speakers might perform similarly, as morphological cues for number and tense exist in English and German. As inflectional cues are not common in Mandarin, L1 Mandarin speakers might struggle the most with number and tense compared to the other two L1 speakers. If transfer is wholesale, then we might expect L1 German speakers to outperform L1 English speakers because of the greater profusion of morphological cues, who in turn would outperform L1 Mandarin speakers. Our study thus aimed to test whether there is an effect of L1 language background on morphology learning and whether there is a wholesale or feature-by-feature transfer.

Current study

In this study, we examined the extent to which adults can learn the meaning of morphological cues with implicit exposure to the language in a CSL paradigm, determining whether learning varies for morphological cues with different functional targets (that vary in opacity) and assessing the extent to which implicit acquisition of morphological cues depends on the speakers' affinity with morphological cues present within their L1. We tested the effect of transparency of different morphological cues within the novel language, in terms of the extent to which each of the morphemes is

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

dependent or independent of syntactic constraints. Selecting tense, number, and SV agreement thus varied the transparency of the cues, i.e., whether there are immediately observable referents to the morphemes within the environment. We thus determined the limits of CSL in supporting participants' acquisition of different morphological targets—tense, number, and SV agreement—at the same time as learning nouns and verbs.

Our CSL paradigm was inspired by Rebuschat et al. (2021), following a similar design in which participants were asked to make decisions on sentence-scene mapping after hearing one sentence in an artificial language and being presented with two scenes per trial. While the learning conditions do not reflect the full complexities of natural language learning, this design enables us to focus on two aspects of natural language immersive learning and determine their impact on language learning. First, our paradigm reflects the need for language and (visual) environmental information to be coordinated for learning. Second, our design enables us to determine how referential ambiguity among multiple morphemes occurring in each utterance and multiple potential referents in the scene around the learner affects processing.

In our study, we controlled cue salience and transparency of the grammatical morphemes of tense and number—these affixes were in CV (consonant + vowel) form and their referents were visually available in every scene. This allowed us to identify the effect of transparency by comparing the learning between these visually available cues (tense and number) and a visually unavailable cue which was SV agreement. This experimental design enables us to hone in on exactly which aspect of the language has been learned by manipulating the presence of information in the visual scene, allowing us to target acquisition of word stems or morphemes. An accompanying grammaticality judgment task (GJT) enabled us to determine whether the SV agreement was acquired. Although morphological cues in natural language can be more abstract and are not always available in the environment, such as whether an event occurred in the past, present, or future, the design of our study tests for the first time whether participants can isolate the morphological segments from continuous speech and keep track of visual referents that consistently co-occur with them. Having two scenes per learning trial also means that we increase the ambiguity of the possible word-referent mappings available in the environment, mimicking that of naturalistic language learning situations (e.g., Yu & Ballard, 2007), and enables us to test learning online, as it proceeds with exposure.

Furthermore, in an exploratory analysis, we also investigated whether explicit knowledge of morphology emerges from the implicit CSL learning environment without any instruction or feedback. Learners may become aware of the meanings of number and tense affixes in our studies, as explicit knowledge has consistently been demonstrated to arise from implicit CSL learning (Ge et al., 2025; Monaghan et al., 2019). Conversely, participants might not recognize the SV agreement since it has been shown that incidental exposure cannot support awareness of functional morphology, such as SV agreement (Kachinske & DeKeyser, 2024). We also expect to see that L1 German speakers exhibit greater awareness of SV agreement than L1 English and Mandarin speakers, given that L1 morphological richness has been evidenced to correlate with the awareness of non-native languages (Wu & Juffs, 2022). However, it remains unclear whether CSL learning can facilitate explicit morphological knowledge and whether this awareness interacts with L1 morphological richness, the result of which gives implications for the types of morphology that require explicit language instruction.

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

Research questions and predictions

Our first research question asked whether grammatical morphemes could be learned from CSL alongside word stem learning. Accurate predictions are currently difficult to ascertain for this research question, as adults were evidenced to be able to learn the abstract grammatical morphemes from cross-situational statistics (Finley, 2023) but learning morphology from sentence-based language input was challenging (Rebuschat et al., 2021).

Our second research question asked whether there are differences in learning morphological features according to their transparency. Learning of tense, number, and SV agreement varied in terms of transparency of the target. We predicted that tense and number would be easier to acquire than SV agreement, as previous research indicated that transparency affects learning (Hofweber et al., 2023; Sehyr & Emmorey, 2019). However, it is difficult to predict whether tense and number would differ from one another, as both of the visual referents to tense and number were made transparent in our experiment.

The third research question investigated the extent to which morphological learning difficulty is affected by L1 background (English, German, Mandarin). We tested whether differences in morphological expressiveness in learners' L1(s) affected the acquisition of different morphological features from cross-situational statistics. We test two theories of transfer by comparing the learning of tense and number affixes between L1 English, German, and Mandarin speakers. If there is a wholesale transfer effect from L1 to additional language learning of morphology, then L1 morphological richness would affect learners' sensitivity toward (all of) the morphological features in the novel language, such that the German group should outperform both the English and Mandarin L1 groups. However, if transfer is feature by feature according to structural similarity (e.g., Westergaard et al., 2017) then the English and German groups should be similar to one another and outperform the Mandarin L1 group. However, the cross-linguistic transfer effect may be limited to semantically transparent, interpretable features, like number and tense in our paradigm. In contrast, participants from different L1s may demonstrate similar accuracy in SV agreement tasks, as its difficulty is likely to be driven by real-time processing issues rather than interference from L1 (Lago et al., 2025).

Finally, as an explorative research question, we also assessed participants for their explicit awareness of the different morphological features and determined whether this explicit knowledge related to language acquisition of different aspects of the language. Based on previous studies (e.g., Ge et al., 2025; Monaghan et al., 2019), we predicted that the awareness of the knowledge would predict the learning of different linguistic features.

The design and analysis were preregistered before data collection (https://osf.io/x6svp).

Methods

Participants

Sample size was estimated using Monte Carlo simulations of data, which predicted that 35 participants per language group would be sufficient for power of .8 to find medium size effects (Cohen's d = .5) of main effects of morphological feature type and language background affecting learning overall. The detailed description of our power analysis can be found with our materials, data, and analysis scripts (https://osf.io/dvpgq/?view_only=5ce6d476492e42ccae7e906f178fdfcc).

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

One hundred and seventeen L1 English, German, and Mandarin native speakers voluntarily participated in this study. However, 10 participants had to be excluded either because they took written notes during the experiment or because their language background did not meet the inclusion criteria of being either an L1 English, German, or Mandarin speaker. Due to technical issues, two participants had to be excluded due to their missing data in the CSL task and debriefing questionnaires. Our final sample thus consisted of 105 participants (74 women, 31 men), which were distributed into three groups, based on their L1(s).

Thirty-five participants each spoke Mandarin, English, and German as their L1(s). None of the participants reported having learned Portuguese, on which the phonetics of the artificial language were based. However, all participants in L1 Mandarin and L1 German groups reported having learned a non-native language that marked number, tense, and SV agreement, mostly English, whereas 66% (23/35) of participants in L1 English group reported being monolingual. The mean age in our sample was 23.27 (standard deviation = 4.72, range = 18–36 years), and there were no significant differences between the groups in terms of age. Participants were recruited via social media (L1 Mandarin), word of mouth (L1 German), or via their institution's participant panel (L1 English).

The study was approved by the ethics review panel of the Faculty of Arts and Social Sciences at Lancaster University and conducted in accordance with the provisions of the World Medical Association Declaration of Helsinki. None of the participants were remunerated in this study. However, the L1 English group received course credits at their home institution for taking part.

Materials

Artificial language

Vocabulary. The artificial language consisted of 16 disyllabic pseudowords: Half functioned as nouns and half as verbs. Nouns referred to eight distinct cartoon animals (e.g., panda, pig), while verbs denoted common actions (e.g., working, walking, sleeping). An additional five monosyllabic pseudowords served as grammatical morphemes marking number (singular/plural) and tense (past, present, future). All vocabulary items were recorded in a monotone by a female native speaker of Portuguese.

The eight cartoon animals served as visual referents for the noun vocabulary. Each was shown performing the eight actions, either alone or two of the same animals, depending on the number. Tense was indicated by visual time cues presented above the action: a written word in the participants' corresponding L1s, paired with an icon (left arrow = past, circle = present, right arrow = future), as illustrated in Figures 1 and 2.[1]

We randomized four lists of word-referent mappings to reduce the impact of a particular mapping being easier to learn. The 16 disyllabic pseudowords were randomly mapped to the characters and the actions, and the five monosyllabic pseudowords to the different grammatical morphemes. Table 1 presents the artificial language vocabulary

[1]Note that in previous studies teaching tense in artificial languages, researchers have trained participants on the meaning of time adverbs (e.g., Ellis, 2007) prior to exposure to the grammar of the language, whereas we wanted to know whether participants could acquire the tense without this prior training, and without reference to adverbs. Note also that time tends to be represented in terms of left (past) and right (future) in all the L1 language groups, though L1 Mandarin speakers also tend to represent time as down (past) and up (future) (Gu, Zheng & Swerts, 2019).

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

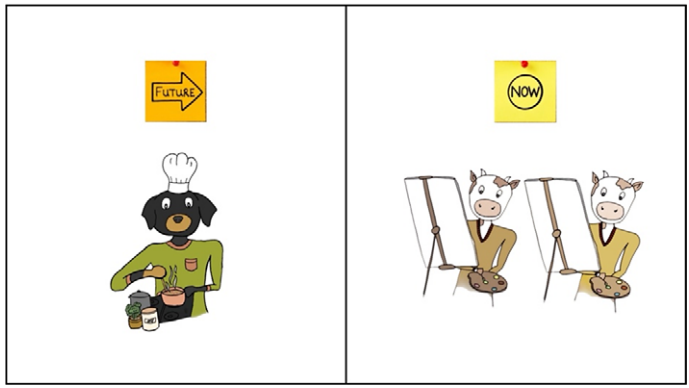


Figure 1. An example of a training trial in the (English) CSL task. Participants were presented with two scenes depicting animal(s) performing an action and played a single artificial language sentence (e.g., /luṯʃisaɪ naɪpəpaʊsaɪ/). Their task was to decide, as quickly and accurately as possible, which scene the sentence referred to. The trial was used with the L1 English group, so the time indicators are presented in English.

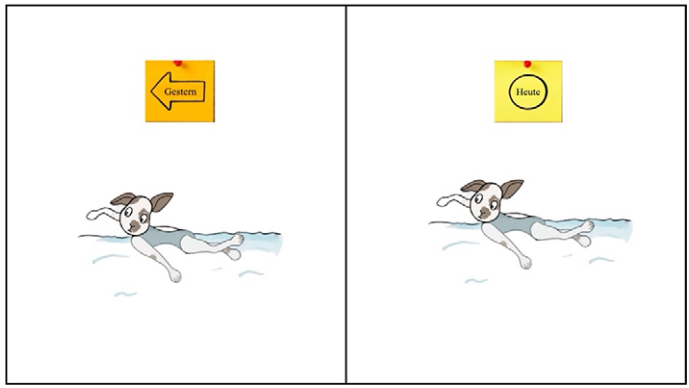


Figure 2. An example of a tense test trial in the (German) CSL task. In test trials, the two scenes were identical with a single difference. In the example trial, only the time of the event is different between the two scenes, as seen in the time indicators "Gestern" (German, yesterday) and "Heute" (today). The trial in this example tests if participants have learned the past tense morphemes as the agent and the action. The trial was used with the L1 German group, so the time indicators are presented in German.

and their respective meanings, for one of the four random mappings. The complete set of random mappings can be found in Table S-1 in supplementary information. The animal cartoon characters can be found in Appendix A. The entire set of images can be found on our Open Science Framework (OSF) site.

Grammar. The artificial language sentences were intransitive and followed SV order. The number morphemes were attached to both the nouns and the verbs, and the tense morphemes were only to the verbs. That is, each noun consisted of the stem plus the number suffix (singular or plural), and each verb consisted of the stem, followed by a

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

tense suffix (past, present, future) and a number suffix (singular or plural). The double marking of number suffixes was employed, which is used to indicate SV agreement in some natural languages such as Spanish. In Spanish, both the subject noun and the verb are morphologically marked for number—for instance, los niños cantan ("the children sing"), where plural suffixes appear on both the noun (-s) and the verb (-an). It is worth noting that in our artificial language, the number was marked using identical suffixes on both subjects and verbs. While this reinforces form-meaning mappings, it may have unintentionally reduced the overall complexity of the language by increasing the transparency and predictability of the morphological system. In natural languages such as English, SV agreement often exhibits lower contingency than in our artificial language. For instance, the morpheme -s marks plural nouns (e.g., dogs) but singular verbs (e.g., walks), creating inconsistencies in form-meaning mappings. Such low contingency between form and meaning has been shown to impede learnability (Ellis, 2022). In contrast, the double marking system in our artificial language featured highly contingent, one-to-one mappings between form and meaning—for example, [ti] unambiguously marked plurality, and [sai] consistently marked singular. This increased transparency represents a limitation of the current study, as it may have made the morphological system easier to acquire than what learners would encounter in some natural languages.

For example, following the first randomization in Table 1, the sentence /faʊluti paʈʃunɑti/ would mean "The pandas worked" and would be constructed as follows:

(1) /faʊluti paʈʃunɑti/
Faulu ti pachu na ti
Panda PL work PAST PL
"The pandas worked."

Table 1. The artificial language vocabulary used in this study

Category		Pseudowords	Meaning
Nouns		/faʊlu/	panda
		/fima/	pig
		/fuki/	lion
		/jitu/	mouse
		/kitə/	sheep
		/lipə/	rabbit
		/luʈʃi/	dog
		/ʃaji/	cow
Verbs		/naɪpə/	cook
		/paʈʃu/	work
		/paʃə/	swim
		/siʃə/	run
		/pulə/	sleep
		/suli/	walk
		/masə/	sing
		/tusi/	paint
Morphemes	Number	/saɪ/	singular
		/ti/	plural
	Tense	/nɑ/	past
		/kə/	present
		/paʊ/	future

Notes: There were eight nouns, eight verbs, and five grammatical morphemes (tense and number marking). There were four random pseudoword-referent mappings to avoid preexisting biases. Here, we report one of the randomizations.

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

We constructed 504 artificial language sentences, comprising 480 sentences in the cross-situational learning (CSL) task (384 in training trials and 96 in test trials) and 24 sentences in the grammaticality judgment task (GJT).

Retrospective verbal reports

We used a questionnaire to gather retrospective verbal reports (Rebuschat, 2013). These allowed us to determine if participants became aware of (aspects of) the artificial language, and if so, which ones (Rebuschat, 2013). The questionnaire was adapted from Rebuschat et al. (2015). We started by asking participants to report any strategies they might have used during the CSL task. Specifically, we asked how they decided which scene was the correct referent of the sentence, whether they were just guessing or whether strategies had been applied. A follow-up question regarding the strategies was asked that if their strategies had changed throughout the CSL task. In the second section, we investigated the degree of awareness regarding the meaning of the grammatical morphemes. The questions gradually prompted them with more and more explicit information. We started with a general question by asking whether they noticed any patterns or rules about the grammatical structure. This is followed by more specific questions asking whether they had noticed and realized the meaning of the sound segments *na*, *ke*, *pau*, *sai*, and *ti*. Finally, we asked them what they thought the aim of the study was. The questionnaire can be found in Appendix B.

Procedure

After providing informed consent, participants completed the language background questionnaire, followed by two tasks: CSL and grammaticality judgments (GJT), as described below. Finally, they completed the debriefing questionnaire (retrospective verbal reports). The entire procedure took around 60 min.

Cross-situational learning task

The CSL task was used to train and test participants on the acquisition of novel nouns, verbs, and morphemes. Participants were informed that they would hear an artificial language sentence and see two scenes on the screen. Their task was to decide, as quickly and accurately as possible, which scene the sentence referred to. No feedback was provided.

During each trial, participants were first presented with a fixation cross for 500 ms, followed immediately by the presentation of two static scenes, one on the left and one on the right side of the screen; 1,000 ms later, participants were then played an artificial language sentence describing one of the two scenes. Immediately after the sentence finished playing, participants had to indicate, as quickly and as accurately as possible, which scene the sentence referred to. They were instructed to press Q on the keyboard for the left scene or P for the right scene. No time limits were set for each trial; the next trial was only played after participants entered a response. Figure 1 provides an example of a training trial of the CSL task.

There were two types of CSL trials, training trials and test trials, which differed in terms of the number of elements that vary between the correct and foil scenes. In the training trials, to reflect ambiguity of potential referents in the environment, the target and the foil scenes differed in terms of two to four elements (different agents, actions, number, and/or time cues); for example in Figure 1, the scenes differed in agent, action,

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

number, and time. However, in the test trials, the two scenes only differed in one aspect. This manipulation allowed us to test what nouns, verbs, and grammatical morphemes participants had learned as they completed the CSL task. To test noun learning, the two scenes were identical except for the agents (same actions, same time cues, different agents). To test verb learning, the two scenes were identical except for the actions (same agents, same time cues, different actions). To test morpheme learning, the two scenes were identical except for the time cues (same agents, same actions, different time cues) or they were identical except for the number of agents (same actions, same time cues, but different number of agents). Figure 2 provides an example of a test trial for tense—the correct scene can be selected only if the tense morpheme is known.

The CSL task consisted of eight blocks, each of which contained 48 training trials. We carefully balanced the presentation frequency of nouns, verbs, and morphemes across blocks; the frequency of target and foil scenes; and their respective locations on the screen (left or right). Importantly, blocks 4, 6, and 8 of the CSL task also contained 32 test trials, in addition to the 48 training trials. That is, in these blocks, both training and test trials occurred in random sequence. Of the 32 test trials, eight were designed to test noun learning, eight to test verb learning, eight to test the acquisition of the number morphemes, and a further eight to test the acquisition of the tense morphemes. The advantage of mixing training and test trials in specific blocks was that it enabled us to test what aspects of the language participants were acquired first (see Rebuschat et al., 2021).

Nouns, verbs, and morphemes occurred an equal number of times within each block, except for one morpheme for tense (past/now/future) that unavoidably occurred one time less than the others, but the differences were even over three blocks. Animal, action, tense, and number features in the pictures occurred an equal number of times in both target and foil scenes, except for time indicators for the same reason. Since no feedback was given, if participants were able to use the cross-situational statistics, then they should be able to learn to distinguish the target from foil scenes by tracking the co-occurrences between particular morphemes and particular features of the scene that always co-occurred. Note that tense and number morphemes occurred in different frequencies over trials—as tense related to three targets (past, present, future) and number related to two (singular, plural). Thus, there was a difference in frequency, which can affect non-native language learning (Ellis, 2002, 2012). We return to consider this point in the Discussion.

Grammaticality judgment task

The GJT was used to test SV agreement. In each trial, participants were first played an artificial language sentence (without any scenes on the screen). After the sentence finished playing, they saw a question mark. Participants were instructed to decide, as quickly and accurately as possible, if the sentence sounded "good" or "bad" to them (in relation to the previously heard artificial language stimuli). They pressed Q if the sentence sounded good to them and P if it sounded bad. There were 24 trials. Half the sentences were grammatical, i.e. they had correct SV agreement, and the other half ungrammatical. The ungrammatical sentences contain mismatching of numerals between subject and verb (e.g., subject singular and verb plural, or subject plural and verb singular). The GJTs occurred three times, always after a mixed CSL block, i.e. after blocks 4, 6, and 8. In each GJT, we presented eight SV agreement test trials (four grammatical, four ungrammatical). The artificial language sentences can be found on our OSF site in the spreadsheets showing the data.

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

Statistical analysis

One sample *t* tests were conducted to identify when accuracies were above chance, allowing us to compare our data to previous studies. We applied the Holm-Bonferroni method to correct for multiple comparisons, which takes into account the ordering of tests to minimize Type I and Type II errors. We predicted that learning would be most likely to exceed chance later in the training, so we ranked the final block first and initial block last for the correction of *p* values. We then used logistic mixed-effects models (Jaeger, 2008) to test our four research questions. The first model investigated predictors that might influence the training trial accuracies, where accuracy was coded as a binary dependent variable. We started from the null model, which included intercepts for random effects of subjects and items (where item was the sentence participants heard) as well as by-subject random slope for block and by-item random slopes for block, L1, and their interaction.[2] To find the best fitting model, we added fixed effects of block (1 to 8), L1 (Mandarin, English, and German), and block:L1 and tested the improvement in model fit using log-likelihood tests. We also tested the quadratic effect for block after other fixed effects were entered, to determine whether learning was linear or varied over time.

The next two models investigated factors that might influence the accuracies in the morphology test trials and the nouns and verbs tests trials. For the analysis of the morphological features, the random intercepts in the null model were subjects and items as well as by-subject slopes for block, morphology test types, and by-item random slopes for block and L1.[3] For the noun and verbs tests, the model was exploratory because we did not preregister this analysis. We started with a null model, which included the random intercept of subjects and items as well as by-subject slopes for block, noun versus verb, and by-item random slopes for block and for L1. The final exploratory mixed effects model investigated whether adding awareness (aware versus unaware), Awareness:block, and awareness:L1 improved model fit for the second mixed effects model.

Results

Learning during the cross-situational learning task

The performance on the training trials of the CSL task is displayed in Figure 3, and Table S-2 in supplementary information provides a detailed summary of the performance on the training trials by block and by language group.

The first mixed-effects model investigated the predictors that have effects on training trials' accuracy, to address the first research question about whether CSL is sufficiently powerful to drive grammatical morpheme learning alongside word stem learning. An effect of overall learning, with improvement with exposure, would provide preliminary evidence that learning the language was possible. An effect of language group would also begin to address the third research question determining whether language background influenced acquisition of the novel language.

[2]Note that the by items random slopes of block and L1 were not included in the pre-registered model, but we conservatively included these as well to ensure that our model had maximal random effects structure (Barr et al., 2013).

[3]We omitted to include by item random slope of L1 in the pre-registration; our approach to include it is a more conservative approach to testing the effects (Barr et al., 2013).

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

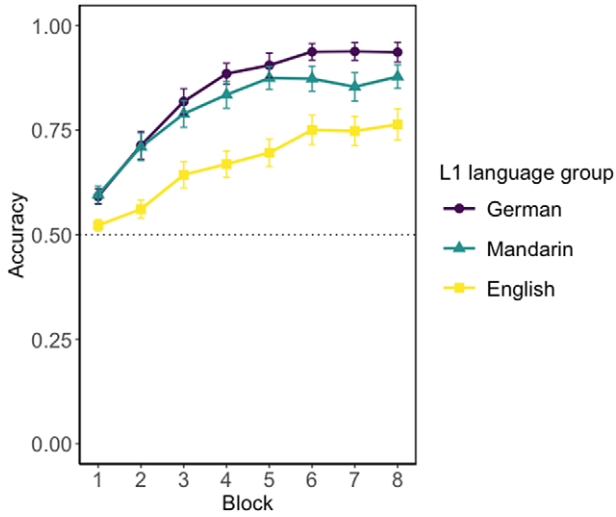


Figure 3. Mean accuracy on the training trials of the CSL task. Error bars represent 95% confidence intervals. The dotted line (.5) shows chance performance.

Beginning with the baseline model containing only random effects, we found that adding the fixed effect of block significantly improved model fit ($\chi^2[1] = 104.19$, $p < .001$), supporting the first research question showing that learning was possible overall (see Figure 3). There was also a significant effect of adding L1 group ($\chi^2[2] = 6.7795$, $p = .034$), indicating that language background did affect acquisition. This was nuanced by the interaction between block and L1 group ($\chi^2[2] = 19.704$, $p < .001$). For the interaction between block and L1, compared to the L1 English group, the L1 Mandarin (logit estimate = .254, standard error [*SE*] = .091, $p = .005$) and L1 German (logit estimate = .422, $SE = .092$, $p < .001$) groups interacted positively significantly with block. However, L1 German compared to the L1 Mandarin group did not interact significantly with block (estimate = .168, $SE = .093$, $p = .072$). The rate of learning over time was thus slower for the L1 English than the L1 German or Mandarin group, showing that language background affected the overall pace of learning, as illustrated in Figure 3. The quadratic effect for block showed no significant difference ($\chi^2[1] = 2.052$, $p = .152$). To ensure there was no bias in word-meaning mappings, we also tested whether different versions of these mappings affected overall learning. The analysis showed that including language version as a predictor did not significantly improve the model fit ($\chi^2[3] = 0$, $p = 1$), indicating no performance differences across word-meaning mappings. The final best-fitting model is reported in Table S-8 in supplementary information.

Distinguishing test types in the cross-situational learning task

The training trials do not tell us what participants' decisions are based on; they could correctly select a scene due to any of the distinctive features it contains. To determine precisely what lexical or morphological features participants have acquired, we analyzed their performance on the test trials of the CSL task. This enables us to test the second research question to uncover the effect of transparency on learning morphemes and

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

to provide further information regarding whether the influence of language background affects learning generally or for particular language features, addressing research question 3.

Figures 4A–D visualize the accuracy of the test trials that occurred in blocks 4, 6, and 8 of the CSL task. Figure 4E visualizes performance across the three groups for GJT, to investigate SV agreement. Tables S-3 and S-4 in supplementary information present detailed analyses of performance by block and language group on each type of test trial.

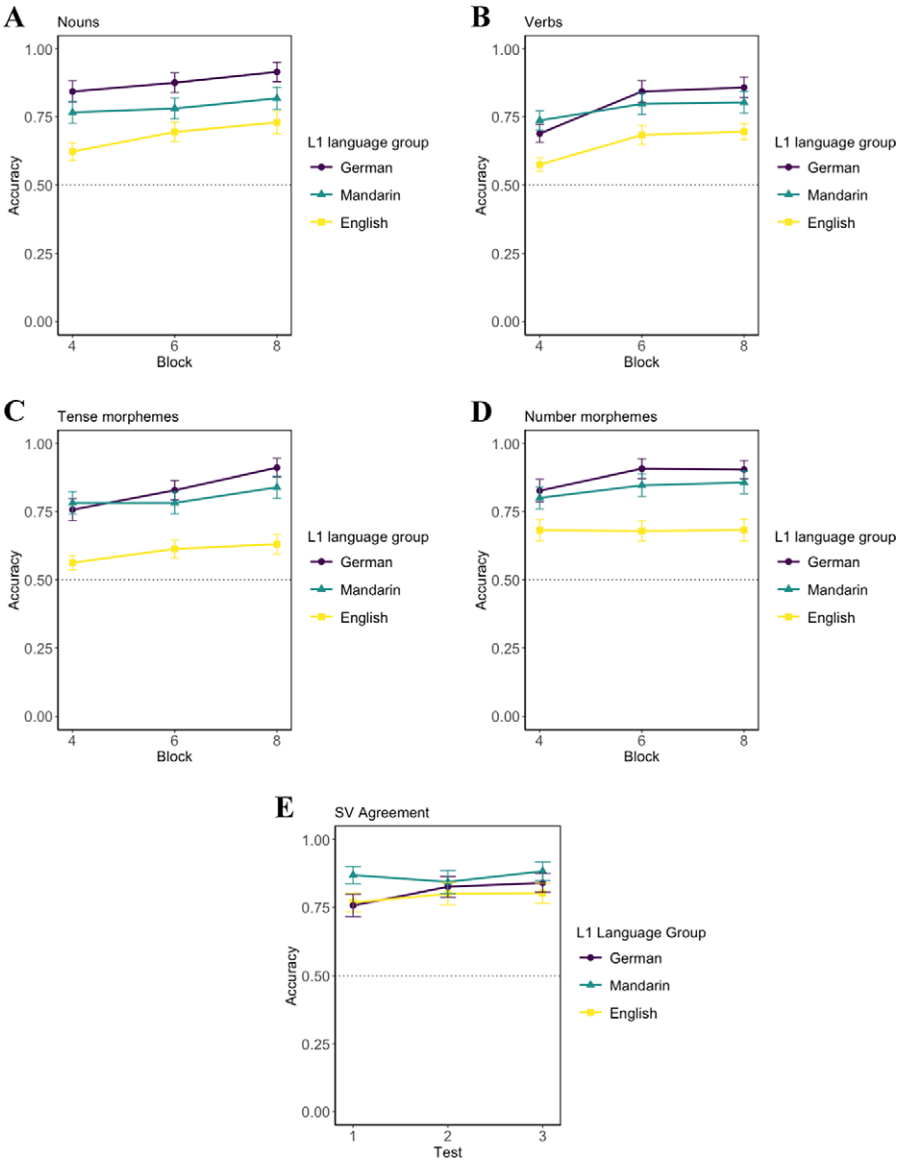


Figure 4. Performance on test trials for lexical categories ([A–D] nouns, verbs, tense, and number morphemes) and syntax ([E] SV agreement). Error bars represent 95% confidence intervals. The dotted line (.5) shows chance performance.

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

To provide descriptive statistics as to whether and when adult learners had acquired each specific feature (nouns, verbs, number morphemes, tense morphemes, SV agreement) via CSL, we ran one-sample *t* tests for the test trials in the three test blocks to determine the time when the performance was greater than chance (with *p* values required to be lower than .05/3 for block 4, lower than .05/2 for block 6, and lower than .05 for block 8, in accordance with Holm-Bonferroni correction). Results for L1 Mandarin, L1 English, and L1 German groups are shown in Tables S-5, S-6, and S-7 in supplementary information. Performance was significantly above chance for all test types in L1 Mandarin and L1 German groups. The L1 English group showed slower learning in tense and verb tests. In L1 English group, the accuracy for tense tests in block 4 ($t[34] = 1.57$, $p = .126$, d = .27) was not significantly above chance yet but it increased to above chance level in block 6 ($t[34] = 2.50$, $p = .017$, d = .42) and block 8 ($t[34] = 2.76$, $p = .009$, d = .47). The accuracy for verb tests in block 4 ($t[34] = 2.08$, $p = .046$, d = .35) was also not significantly above chance until block 6 ($t[34] = 5.35$, $p < .001$, d = .91). This indicates that by block 4, participants in all groups were likely to have acquired all linguistic features in the artificial language, except for the acquisition of tense markers and verbs in L1 English group, the learning of which shows later in block 6 and block 8.

We next tested whether learning varied according to the morpheme type being tested (research question 2) as well as whether the L1 background had a distinct effect on different aspects of the language (research question 3). For this, we conducted a mixed-effects model just on the morphology tests.

For the baseline model, we found that adding block ($\chi^2[1] = 38.829$, $p < .001$) and L1 ($\chi^2[2] = 9.437$, $p = .009$) were significant, showing that learning was possible and progressed with exposure over all test types. There was also a significant effect of morphology test type ($\chi^2[2] = 16.498$, $p < .001$), indicating that accuracies were significantly different among morphology tests. The interaction between morphology test type and L1 group ($\chi^2[4] = 10.269$, $p = .036$) also significantly improved model fit. There was no significant effect of the interaction between morphology test type and block ($\chi^2[2] = 1.056$, $p = .59$).

To interpret the morphology test type by L1 group effect, we conducted post hoc tests comparing language group for each test type and then comparing test types within each language group. For the tense test, L1 Mandarin and L1 German were similar in accuracy (logit estimate = .002, *SE* = .410, $p = .996$) and significantly more accurate than L1 English (logit estimate = –1.275, *SE* = .388, $p = .001$), as shown in Figure 4C. A similar pattern was observed for the number test (Figure 4D). In the SV agreement test, L1 Mandarin was significantly more accurate than L1 German (logit estimate = –1.080, *SE* = .418, $p = .010$). The accuracies in the L1 English group were not significantly different than either the L1 Mandarin ($p = .118$) or German ($p = .274$) groups. For the within-group comparison, in the L1 Mandarin group, we found that number was learned significantly better than tense (logit estimate = .942, *SE* = .344, $p = .006$) but no significant difference between SV agreement and tense (logit estimate = .606, *SE* = .389, $p = .120$) nor between SV agreement and number (logit estimate = –.163, *SE* = .466, $p = .727$). In the L1 English group, SV agreement accuracy was significantly higher than tense (logit estimate = –1.302, *SE* = .353, $p < .001$) but not significantly higher than number (logit estimate = –.497, *SE* = .421, $p = .238$). Number was also significantly higher than tense (logit estimate = .836, *SE* = .288, $p = .004$). In the L1 German group, number was learned significantly better than tense (logit estimate = –.894, *SE* = .359, $p = .013$) and SV agreement (logit estimate = –1.319, *SE* = .464, $p = .005$), while tense was not significantly different from SV agreement tests (logit estimate = –.324, *SE* = .392, $p = .408$). The effect of interaction between L1 and morphology test types suggests that

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

there are distinctions in terms of accuracies on different morphology tests varying between different L1s. The final best-fitting mixed effects model is shown in Table S-9 in supplementary information.

The third mixed-effects model explored whether nouns and verbs were learned differently in the CSL paradigm. We found that adding block ($\chi^2[1] = 43.628$, $p < .001$), L1 ($\chi^2[2] = 13.153$, $p = .001$), and noun versus verb ($\chi^2[1] = 7.557$, $p = .006$) significantly improved model fit over the baseline model containing only random effects. For L1, accuracy in L1 Mandarin (logit estimate = .857, *SE* = .283, $p = .002$) and L1 German (logit estimate = .953, *SE* = .285, $p < .001$) groups were significantly higher than the L1 English group but did not differ from one another (logit estimate = .093, *SE* = .297, $p = .754$). Nouns were learned more readily than verbs, consistent with previous studies (Monaghan et al., 2015); see Figure 4. There were no significant interactions (noun versus verb:L1, $\chi^2[2] = 3.508$, $p = .173$; noun versus verb:block, $\chi^2[1] = .014$, $p = .905$). Hence, L1 background again had an influence on the extent of learning, however, unlike for the grammatical feature morpheme types, this was an overall effect, rather than specific to morpheme type. The best-fitting mixed effects model is shown in Table S-10 in supplementary information.

Retrospective verbal reports

Participants' answers to the debriefing questions were coded according to Rebuschat et al.'s (2015) coding scheme of awareness, ranking from full awareness to complete unawareness (see Appendix B). Participants who reported using morphological rules to distinguish words strategically were considered to have "full awareness" (Q1~2); those who mentioned past, present, future (Q3), singular, or plural (Q4) or specified the SV agreement when asking about the patterns of the language or the morphology system were considered "partial awareness"; and those who only mentioned tense, number, or pattern of sounds were coded as having "minimal awareness." Participants who did not report tense, number, or SV agreement were coded as "unaware." All participants who reported minimal, partial, or full awareness were coded as "aware" and others "unaware."

Following the criteria outlined above, we found that 60 out of 105 participants were fully aware of the morphological rules. Overall, 17 participants in the L1 English group, 30 participants in the Mandarin group and thirty-one participants in the German group were at some level aware of the morphological cues. All participants reported guessing at the beginning, but some later used strategies of calculating the number of categories in the pictures (animals, actions, time, number of animals) to figure out their meaning by comparing similar pictures. And some others reported that they learned from the errors when testing and renewing different assumptions. Seventeen participants at some level noticed the morphological cues. When asked about the meaning of inflectional cues, they either gave a generic answer like number or tense or SV agreement or were more specific about the meaning of each sound (e.g., "Yes, I believe 'sai' is highlighting a single character, while 'ti' stands for many"; "The structure of the sentence is [subject + number morphemes + verb + tense + number morphemes]"). According to the criteria, nine participants were categorized as partially aware of the morphological rule and eight were minimally aware. The rest of the 27 participants reported no awareness of the morphological cues.

Accuracies of CSL tasks between participants who showed different levels of awareness in their debriefing questionnaires are shown in Figure 5. To test the fourth

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

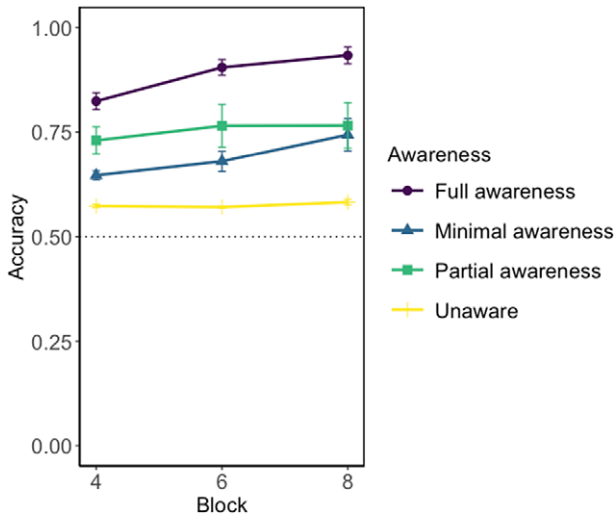


Figure 5. Participants' accuracy on all the CSL tasks, including training and test trials: comparisons between awareness groups (full awareness, partial awareness, minimal Awareness, unaware).
Note: Error bars represent 95% confidence intervals.

research question—determining the role of awareness in learning morphological features in the language—we first conducted descriptive tests using *t* tests on overall performance for each awareness level and then compared aware and unaware for each morphological feature. We then conducted mixed effects models to determine how awareness interacted with different language groups and morphological test types.

Participants who showed full awareness performed significantly better than those with partial awareness ($t[67] = 3.73$, $p < .001$) and minimal awareness groups ($t[66] = 5.420$, $p < .001$). We did not find significant differences between partial-aware and minimal-aware groups ($t[15] = -1.116$, $p = .282$). The partial-aware ($t[34] = 5.197$, $p < .001$) and minimal-aware ($t[34] = 3.511$, $p = .001$) groups showed significantly higher accuracy in CSL tasks than the unaware group. As there was a large and clear discrepancy in CSL performance between people who were aware and unaware of the morphological features, we further explored the performance differences between adults who generate awareness during the immersive learning environment and those who do not. For the following analysis, we included full awareness, partial awareness, and minimal awareness groups in the awareness group, in comparison with the unawareness group, because there was the largest difference in performance between the unaware and minimal awareness groups, and this meant that the results would be consistent with previous studies of awareness and language learning (e.g., Rebuschat et al., 2015).

As shown in Figure 6, participants who showed awareness of tense, number, and SV agreement performed significantly differently from participants who showed unawareness of the related morphological features. Figure 6A shows that participants being aware of the morphological number performed significantly better than participants who were unaware ($t[103] = 8.232$, $p < .001$). In Figure 6B, participants aware of tense are significantly more accurate than the unaware ($t[103] = 6.542$, $p < .001$). Similarly, in Figure 6C, participants being aware of SV agreement performed significantly better than the unaware ($t[103] = 8.085$, $p < .001$).

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

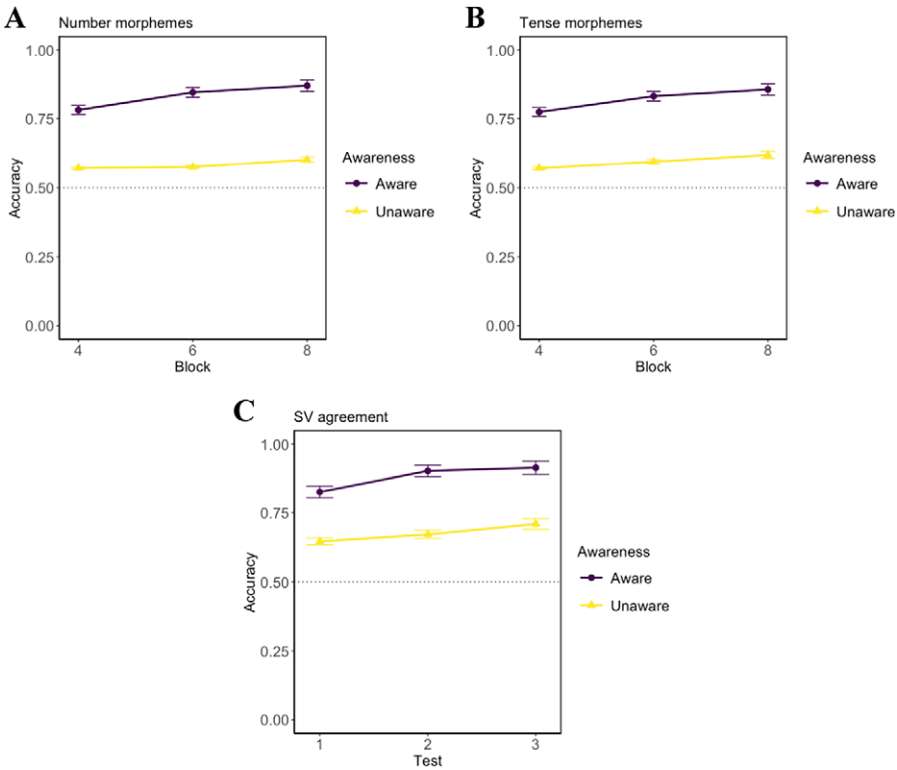


Figure 6. Participants' accuracy on CSL tasks: comparisons between participants who reported being aware and unaware of A number, B tense, and C SV agreement morphological features.
Note: Error bars represent 95% confidence intervals.

For the mixed-effects model, testing the predictors of learning, we found that, as in our second model, above, there were significant effects of block ($\chi^2[1] = 38.853$, $p < .001$), and L1 group ($\chi^2[2] = 7.668$, $p = .020$), as well as Awareness ($\chi^2[1] = 42.738$, $p < .001$) and awareness:block ($\chi^2[1] = 30.159$, $p < .001$). The inclusion of Awareness:L1 did not significantly improve the model fit (χ^2 [2] = 1.798, $p = .407$), which is excluded from the final model. The best-fitting model is detailed in Table S-11 in supplementary information. Taken together, these results showed that awareness predicted the accuracy of the CSL tasks—participants could develop explicit knowledge of the morphological properties of the language and those that did perform more accurately in learning the language. Furthermore, this effect appeared to be consistent across participants from different language backgrounds. The analysis of the debriefing questionnaire demonstrates the importance of awareness of the morphological features in the early acquisition of non-native languages, specifically in an immersive learning setting.

Discussion

In this study, we investigated the extent to which adults can learn the meaning of morphological cues in an implicit learning environment without any explicit

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

instruction of the language. The CSL paradigm was adopted, which mirrors a key aspect of the natural language environment where language and environmental cues co-occur and require coordination but are presented in a controlled laboratory setting. While being exposed to the audio-visual co-occurring events, participants were asked to identify which of two scenes a spoken sentence was referring to. This reflected the ambiguity of reference in a naturalistic language learning environment and enabled us to test how this information can be combined during learning. Importantly, no instruction about the language structure was given, and no feedback was provided throughout the experiment. To investigate whether low transparency makes morphology difficult to learn, we compared the learning of number and tense with S-V agreement of an artificial language. This study also explores the influence of L1 morphological background on the acquisition of non-native languages. We recruited L1 Mandarin, English, and German speakers, ranging from morphologically poor to rich languages.

Our first research question investigated whether grammatical morphemes can be learned from the CSL paradigm alongside word stem. Whereas learning of words is well established (Dal Ben et al., 2023; Monaghan et al., 2015), acquisition of morphology from implicit learning situations, such as CSL, is less often assessed (Finley, 2023; Rebuschat et al., 2021) and never directly compared to word stem learning. In our study, overall robust learning was found even by the first 5 min of exposure for all three L1 groups (L1 English, L1 German, L1 Mandarin); furthermore, participants' performance on both training and test trials significantly improved over time. Our findings from the training trials performance suggest that it is likely adults possess the cognitive ability to rapidly learn both the word stems and affixes at the same time, without any feedback or explicit instruction of their meanings. In the test trials, we found that both word stems and grammatical morphemes, when assessed separately, were acquired from cross-situational statistics. However, word stems were also acquired more accurately than grammatical morphemes, consistent with previous studies of morphological acquisition (Rebuschat et al., 2021). While the results did not reveal a clear pattern indicating that word stems and grammatical morphemes were acquired at different rates in this study, the measures of learning against chance (Tables S-5 to S-7 in supplementary information) did show that learning of word stems may have preceded learning of tense morphemes for the L1 English group. Whereas, by block 4, learning of nouns, verbs, number morphemes, and SV agreement was significantly better than chance, this was not yet the case for tense, which was significantly acquired by block 6. There was no evidence of a difference in precedence for learning word stems over grammatical morphemes for L1 German or L1 Mandarin groups. Thus, for L1 English, the acquisition of tense morphemes may have depended upon prior acquisition of the word stems. For the other language groups, perhaps because learning was overall more accurate, different rates of learning were not apparent.

Our study shows much higher learning of functional markers than previous studies of morphological acquisition (e.g., Finley, 2023; Rebuschat et al., 2021). In Rebuschat et al. (2021), for instance, case markers for subject and object roles were found to be difficult to learn (Monaghan et al., 2021; Rebuschat et al., 2021; Walker et al., 2020). According to Ellis (2022), three key characteristics of morphology that pose learning challenges are salience, contingency, and redundancy. However, compared with previous studies of case marker learning from CSL (Monaghan et al., 2021; Rebuschat et al., 2021; Walker et al., 2020), our morphological cues were similarly salient (in CV form), and yet even with this relatively low salience, learning was still very effective in our study. Therefore, one of the possible reasons for different learning outcomes in our

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

study compared to previous studies of morphological acquisition might result from the transparency of the cue indication in the visual scene (DeKeyser, 2005). Compared to the case marker, the meaning of which needs to be deduced from understanding the relationship between the subject and object in the visual scene, the indication of number and tense in our study is more straightforward. However, learners can rapidly and successfully learn form-meaning mapping even when the cues are more abstract (e.g., Finley, 2023).

In our second research question, we investigated whether we would see differences in learning morphological features according to their transparency. According to the bottleneck hypothesis (Slabakova, 2014), functional morphology, such as SV agreement, is predicted to be the most difficult part to learn in non-native language acquisition, due to the requirement to align the syntactic relations among words in the sentence with properties of the environment. In contrast, other morphological features, such as tense and number, can relate to more transparent features of the environment that do not require inter-relations within the syntax to be simultaneously processed with the environment to which the sentence refers. The results in our study show otherwise. Robust learning was also found in the SV agreement tests in all three L1 groups, which indicates that learners are able to resolve both morphology and syntax. Furthermore, when considering all the language groups together, number was found to be learned significantly more accurately than tense. Frequency might be one of the explanations for the better performance in learning number cues, as each of the number cues had a higher input frequency than either of the tense cues. However, participants in L1 Mandarin and English groups performed significantly better in SV agreement than other linguistic features, even when it appeared less frequent than number cues. Our results did not align with the bottleneck hypothesis, which suggested the learning difficulty of functional morphology comes from the intertwining of different linguistic features (e.g., morphology and syntax). Note that the feature of SV agreement in our artificial language was expressed in the way that the subject and verb have the same ending of number cue, while in English, for example, have the same ending of subject and verb but the number of subjects could vary. For instance, when "-s" appears at the end of the subject, the number of the subject is more than one in the visual scene, but when "-s" appears at the end of the verb, it indicates the singular number of subjects instead. Therefore, our results suggest that the inconsistency of form-meaning mapping, or "low contingency" in Ellis (2022), accompanied by the SV agreement might be the reason for the bottleneck of learning rather than the intertwining of linguistic features. Our results also found that tense and SV agreement, however, were not significantly different. There are several possible reasons for this pattern. First, it may be that the differences in the transparency of the target varied. Number may have been more apparent in the visual scenes than tense or SV agreement. However, tense was more apparent than SV agreement, which had to be deduced from the relations among stimuli in the language, and with the exception of the L1 German group, SV agreement was learned more easily overall. Alternatively, tense may have been more difficult to acquire, not because of transparency but because it is less frequent than SV agreement as well. As tense varied over three targets rather than two (for number and SV agreement). This could have affected learning—with lower performance on tense due not to relative difficulties in the acquisition of tense but rather due to difficulties in learning a three-way rather than a two-way morphological system. Varying tense across just two temporal states—e.g., past and present—will enable us to distinguish learning effects associated with frequency and variability from those relating to transparency.

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

Our third research question addressed whether L1 background influences ease of learning of different aspects of morphology from CSL and also whether there were knock-on effects for word stem learning. In our results, we found that for the training trials, German and Mandarin speakers learned more quickly than did the English participants. When this was investigated in terms of the individual morphological features, we found that having a morphologically rich L1 (German) did not have an overall benefit over a more morphologically impoverished L1 (Mandarin). However, the L1 German group performed significantly better in number and tense tests compared with L1 English group, suggesting a wholesale transfer effect, rather than a feature-by-feature transfer effect, which is consistent with previous findings (e.g., van der Slik et al., 2019). Results in the SV agreement tests also seem to support the claim that L1 transfer only applies to the interpretable linguistic features such as tense and number in our studies but not SV agreement (Lago et al., 2025).

What can explain the overall higher performance in the L1 Mandarin group compared to the L1 English participants? The lower performance in the L1 English group could be because a majority of L1 English speakers in our study were monolingual, potentially a disadvantage compared to the other two groups where multilingualism was prevalent (Nation & McLaughlin, 1986). The effective and accurate learning of the morphological markers in our study, particularly for the L1 Mandarin and German groups, might be due to the fact that most L1 Mandarin speakers and L1 German speakers reported having a high proficiency in at least another language, which might improve their learning of morphology because of prior experience of explicit learning of numbers and tense markers in the non-native language classroom (Bono, 2011). Although L1 Mandarin speakers do not have number and tense markers in their L1, the experience of explicitly learning another language might have turned their attention to those markers during the CSL (Thomas, 1988).

Another possible explanation for good performance on tense markers of the L1 Mandarin group is that they were able to link the morphological features in the current study with time adverbs present in Mandarin. While inflectional morphological cues do not exist in Mandarin, time indications like 了*(le)* and 过*(guo)* are often in CV form. From our study, it is not possible to distinguish categorically whether participants were learning morphological inflections or treating the linguistic indicators of past, present, and future as adverbs. Ellis and Sagarra's (2011) study of tense learning, for instance, implemented time with a similar referent in the visual scene to which the language related but tested also the role of both an adverb and an inflection indicating tense. In that study, again it is not possible to ensure that the inflection was processed as a suffix inflection rather than as another adverb, but the fact that L1 Mandarin participants were poorer at acquiring the inflection but not the adverb (due to the inflection being at odds with the structure of the learners' L1s), suggested that such a suffix operated in a similar way to an inflection for the learners. Thus, a limitation of the current study lies in the potential language background and demographic differences other than morphological richness between L1 groups that could account for the learning difference.

For all L1 groups, tense was learned with less accuracy than number; however, relative ease of learning of SV agreement varied by L1 group. This was easiest to acquire for L1 Mandarin and English groups but more difficult for the L1 German group. This could be due to a negative transfer from German L1, as similar grammatical agreements largely exist in the German language—for example, between adjective and object depending on the gender and number as well as the case. For English, SV agreement only exists in the simple present tense, possibly resulting in no L1 transfer effects being observed.

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

There may be additional demographic or motivational differences between the groups that also resulted in different performance. For example, groups differed in how they were recruited (word of mouth, social media, or participant pools), as outlined above. Measuring additional demographic or motivation characteristics of learners may well help us pinpoint where overall effects in learning come from. Nevertheless, these overall effects of L1 group on learning cannot readily explain the more subtle interaction between L1 group and morphological feature—general background properties of participants are very unlikely to result in better performance only for certain morphemes. These interactions are more likely to be a consequence of language background exerting its effect on learning. Note that the effect of L1 background on learning provides support for the validity of use of a laboratory-based CSL study of language learning—if participants were merely approaching the task as a problem-solving exercise, then we would be unlikely to find effects of different language background on combining the language and environmental information in the task.

Finally, in exploratory analyses, we determined whether participants might be able to acquire explicit knowledge of the language structure and if that was related to their learning. We found that a considerable number of participants were able to report the linguistic structure relevant to morphology as a consequence of implicit statistical exposure to the language from CSL, which is aligned with findings in previous studies that explicit knowledge can arise from implicit exposure (e.g., Ge et al., 2025; Monaghan et al., 2019). However, we did not find significant differences in awareness between groups, which did not confirm the previous finding that L1 morphological richness correlates with awareness (Wu & Juffs, 2022). Moreover, our findings align with theoretical perspectives that conceptualize implicit and explicit learning as interrelated and dynamic processes (e.g., Rebuschat, 2013; Ellis, 2005). Specifically, the results suggest that awareness can emerge from implicit learning, particularly under conditions where the input is statistically rich, well-structured, and consistent (Rebuschat, 2013). In the context of non-native language acquisition, this implies that explicit instruction aimed at drawing learners' attention to morphological features may enhance subsequent implicit learning of the language.

Conclusion

Our study provides the first evidence that learners can acquire multiple morphological features of a language simultaneously, accurately, and in tandem with learning word stems from a language via CSL, challenging the notion that adults struggle with morphology acquisition primarily due to biological constraints.

The current study also provides insight into which aspects of morphology might influence its acquisition. We showed that the morphological features of a learners' L1 affected acquisition, sometimes in surprising ways, with greater learning of morphological properties that were less fully expressed in L1.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S027226312510106X.

Data availability. Our power analysis, materials, anonymized data, and data analysis scripts are available on our project site (https://osf.io/dvpgq/?view_only=5ce6d476492e42ccae7e906f178fdfcc) on an OSF platform.

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

Acknowledgments. We would like to thank Alina Hafki for creating the animal characters used in the experiment and Susana Correia for her support in developing the pseudowords and for recording the audio stimuli. We are grateful to our participants for volunteering their time. Finally, we would like to acknowledge the financial support provided by Lancaster University's Camões Institute Chair (*Cátedra*) for Multilingualism and Diversity.

Competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author note. Padraic Monaghan and Patrick Rebuschat jointly supervised this project and thus shared senior authorship.

References

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.

Bono, M. (2011). Chapter 3: Crosslinguistic interaction and metalinguistic awareness in third language acquisition. In G. De Angelis & J.-M. Dewaele (Eds.), *New trends in crosslinguistic influence and multilingualism research* (pp. 25–52). Multilingual Matters. https://doi.org/10.21832/9781847694430-004

Choi, S. H., & Ionin, T. (2021). Plural marking in the second language: Atomicity, definiteness, and transfer. *Applied Psycholinguistics*, *42*, 549–578. https://doi.org/10.1017/S0142716420000569

Cintrón-Valentín, M. C., & Ellis, N. C. (2016). Salience in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01284

Dal Ben, R., Prequero, I. T., Souza, D. de H., & Hay, J. F. (2023). Speech segmentation and cross-situational word learning in parallel. *Open Mind*, *7*, 510–533. https://doi.org/10.1162/opmi_a_00095

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? Review of issues. *Language Learning*, *55*, 1–25. https://doi.org/10.1111/j.0023-8333.2005.00294.x

Diaubalick, T., & Guijarro-Fuentes, P. (2019). The strength of L1 effects on tense and aspect: How German learners of L2 Spanish deal with acquisitional problems. *Language Acquisition*, *26*, 282–301. https://doi.org/10.1080/10489223.2018.1554663

Ellis, N. C. (2002). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition*, *24*, 297–339. https://doi.org/10.1017/S0272263102002140

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, *27*, 305–352.

Ellis, N. C. (2007). Blocking and Learned Attention in Language Acquisition. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *29*(29).

Ellis, N. C. (2012). What can we count in language, and what counts in language acquisition, cognition, and use? In S. Th. Gries & D. Divjak (Eds.), *Frequency effects in language learning and processing* (pp. 7–34). Walter de Gruyter. https://doi.org/10.1515/9783110274059.7

Ellis, N. C. (2022). Second language learning of morphology. *Journal of the European Second Language Association*, *6*, 34–59. https://doi.org/10.22599/jesla.85

Ellis, N. C., & Sagarra, N. (2011). Learned attention in adult language acquisition: A replication and generalization study and meta-analysis. *Studies in Second Language Acquisition*, *33*, 589–624.

Finley, S. (2023). Morphological cues as an aid to word learning: A cross-situational word learning study. *Journal of Cognitive Psychology*, *35*, 1–21. https://doi.org/10.1080/20445911.2022.2113087

Finn, A. S., & Hudson Kam, C. L. (2015). Why segmentation matters: Experience-driven segmentation errors impair "morpheme" learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1560–1569. https://doi.org/10.1037/xlm0000114

Gardner, Q., Branigan, H. P., & Chondrogianni, V. (2021). Spoken and written production of inflectional morphology among L1 Mandarin speakers of English. *Journal of Memory and Language*, *120*, 104250. https://doi.org/10.1016/j.jml.2021.104250

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

Gass, S. M., Spinner, P., & Behney, J. (Eds.). (2018). *Salience in second language acquisition*. Routledge, Taylor & Francis Group.

Ge, Y., Monaghan, P., & Rebuschat, P. (2025). The role of phonology in non-native word learning: Evidence from cross-situational statistical learning. *Bilingualism: Language and Cognition*, *2*, 15–30. doi:10.1017/S1366728923000986

Gu, Y., Zheng, Y., & Swerts, M. (2019). Having a different *pointing* of view about the future: The effect of signs on co-speech gestures about time in Mandarin-CSL bimodal bilinguals. *Bilingualism: Language and Cognition*, *22*, 836–847.

Hawkins, R., & Hattori, H. (2006). Interpretation of English multiple wh-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research*, *22*, 269–301. https://doi.org/10.1191/0267658306sr269oa

Hofweber, J., Aumônier, L., Janke, V., Gullberg, M., & Marshall, C. (2023). Which aspects of visual motivation aid the implicit learning of signs at first exposure? *Language Learning*.

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

Kachinske, I., & DeKeyser, R. (2024). The role of awareness in implicit and explicit knowledge. *International Review of Applied Linguistics in Language Teaching*, *63*, 857–891. https://doi.org/10.1515/iral-2022-0212

Lago, S., Oltrogge, E., Stone, K., Lago, S., Oltrogge, E., & Stone, K. (2025). Does productive agreement morphology increase sensitivity to agreement in a second language? *Glossa Psycholinguistics*, *4*.

Monaghan, P., & Mattock, K. (2012). Integrating constraints for learning word–referent mappings. *Cognition*, *123*, 133–143. https://doi.org/10.1016/j.cognition.2011.12.010

Monaghan, P., Mattock, K., Davies, R. A. I., & Smith, A. C. (2015). Gavagai is as gavagai does: Learning nouns and verbs from cross-situational statistics. *Cognitive Science*, *39*, 1099–1112. https://doi.org/10.1111/cogs.12186

Monaghan, P., Ruiz, S., & Rebuschat, P. (2021). The role of feedback and instruction on the cross-situational learning of vocabulary and morphosyntax: Mixed effects models reveal local and global effects on acquisition. *Second Language Research*, *37*, 261–289. https://doi.org/10.1177/0267658320927741

Monaghan, P., Schoetensack, C., & Rebuschat, P. (2019). A single paradigm for implicit and statistical learning. *Topics in Cognitive Science*, *11*, 536–554. https://doi.org/10.1111/tops.12439

Nixon, J. S. (2020). Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition*, *197*, 104081. https://doi.org/10.1016/j.cognition.2019.104081

Nation, R., & McLaughlin, B. (1986). Novices and experts: An information processing approach to the "good language learner" *problem.* Applied psycholinguistics, *7*, 41–55.

Quine, W. V. (1960). *Word and object*. Technology Press of the Massachusetts Institute of Technology.

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, *63*, 595–626.

Rebuschat, P. (2022). Implicit learning and language acquisition: Three approaches, one phenomenon. In A. S. Reber & R. Allen (Eds.) *The cognitive unconscious: The first half-century*. Oxford University Press.

Rebuschat, P., Hamrick, P., Riestenberg, K., Sachs, R., & Ziegler, N. (2015). Triangulating measures of awareness: A contribution to the debate on learning without awareness. *Studies in Second Language Acquisition*, *37*, 299–334.

Rebuschat, P., Monaghan, P., & Schoetensack, C. (2021). Learning vocabulary and grammar from cross-situational statistics. *Cognition*, *206*, 104475. https://doi.org/10.1016/j.cognition.2020.104475

Schroer, S. E., & Yu, C. (2023). Looking is not enough: Multimodal attention supports the real-time learning of new words. *Developmental Science*, *26*, e13290. https://doi.org/10.1111/desc.13290

Scott, R. M., & Fisher, C. (2012). 2.5-Year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, *122*, 163–180. https://doi.org/10.1016/j.cognition.2011.10.010

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

Sehyr, Z. S., & Emmorey, K. (2019). The perceived mapping between form and meaning in American Sign Language depends on linguistic knowledge and task: Evidence from iconicity and transparency judgments. *Language and Cognition*, *11*, 208–234. https://doi.org/10.1017/langcog.2019.18

Slabakova, R. (2014). The bottleneck of second language acquisition. *Foreign Language Teaching and Research*, *46*, Article 4. https://eprints.soton.ac.uk/367089/

Slabakova, R. (2017). The scalpel model of third language acquisition. *International Journal of Bilingualism*, *21*, 651–665. https://doi.org/10.1177/1367006916655413

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Suanda, S. H., & Namy, L. L. (2012). Detailed behavioral analysis as a window into cross-situational word learning. *Cognitive Science*, *36*, 545–559. https://doi.org/10.1111/j.1551-6709.2011.01218.x

Suethanapornkul, S. (2020). *Statistical learning of predictive dependencies in the tense-aspect system of a miniature language by English and Thai first language adults* [Doctoral dissertation]. Georgetown University. https://www.proquest.com/docview/2446728798/abstract/ED9C206917AE4F90PQ/1

Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. *Journal of Multilingual and Multicultural Development*, *9*, 235–246. https://doi.org/10.1080/01434632.1988.9994334

Tünnermann, J., & Scharlau, I. (2018). Stuck on a plateau? A model-based approach to fundamental issues in visual temporal-order judgments. *Vision*, *2*, 29. https://doi.org/10.3390/vision2030029

van der Slik, F., van Hout, R., & Schepens, J. (2019). The role of morphological complexity in predicting the learnability of an additional language: The case of La (additional language) Dutch. *Second Language Research*, *35*, 47–70. https://www-jstor-org.ezproxy.lancs.ac.uk/stable/26965674

van Zoest, W., & Donk, M. (2005). The effects of salience on saccadic target selection. *Visual Cognition*, *12*, 353–375. https://doi.org/10.1080/13506280444000229

Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*, 700–711. https://doi.org/10.1037/xlm0000611

Walker, N., Monaghan, P., Schoetensack, C., & Rebuschat, P. (2020). Distinctions in the acquisition of vocabulary and grammar: An individual differences approach. *Language Learning*, *70*, 221–254. https://doi.org/10.1111/lang.12395

Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism*, *21*, 666–682. https://doi.org/10.1177/1367006916648859

Williams, J. N. & Rebuschat, P. (2023). Implicit learning and SLA: A cognitive psychology perspective. In A. Godfroid & H. Hopp (Eds.), *The Routledge handbook of second language acquisition and psycholinguistics* (pp. 281–293). Routledge.

Wu, Z., & Juffs, A. (2022). Effects of L1 morphological type on L2 morphological awareness. *Second Language Research*, *38*, 787–812. https://doi.org/10.1177/0267658321996417

Yu, C., & Ballard, D. H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, *70*, 2149–2165. https://doi.org/10.1016/j.neucom.2006.01.034

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*, 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Yu, C., Zhang, Y., Slone, L. K., & Smith, L. B. (2021). The infant's view redefines the problem of referential uncertainty in early word learning. *Proceedings of the National Academy of Sciences of the United States of America*, *118*, e2107019118. https://doi.org/10.1073/pnas.2107019118

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

Appendix A

The animal cartoon characters used in the CSL task. The entire set of images can be found on our OSF site (https://osf.io/dvpgq/?view_only=5ce6d476492e42ccae7e906f178fdfcc).



Appendix B

The debriefing questionnaire used in the study to elicit retrospective verbal reports.

Q1: How did you decide which picture was the correct referent? Did you just guess throughout the experiment, or did you follow any particular strategies? If so, what strategies did you follow?
Q2: Did you notice any particular patterns or rules about the grammatical structure of this new language?
Q3: Did you notice the sounds "na," "ke," or "pau"? If so, what do you think they mean? (The tense morphemes asked in Q3 variants between different versions of mappings.)

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press

Q4: Did you notice the sounds "sai" or "ti"? If so, what do you think they mean? (The number morphemes asked in Q4 variants between different versions of mappings.)
Q5: Do you think the way you made decisions on the pictures changed throughout the experiment?
Q6: What do you think was the aim of this study?
Note: The questionnaire was adapted from Rebuschat et al. (2015).

Cite this article: Zhu, L., Rebuschat, P., Nixon, J. S., & Monaghan, P. (2025). Learning morphology from cross-situational statistics. *Studies in Second Language Acquisition,* 47: 1250–1276. https://doi.org/10.1017/S027226312510106X

https://doi.org/10.1017/S027226312510106X Published online by Cambridge University Press