Can sign-naïve adults learn about the phonological regularities of an unfamiliar sign language from minimal exposure?

Julia Hofweber
University College London, UK; Northeastern University London, UK

Lizzy Aumônier
University of Kent, UK; Northeastern University London, UK

Vikki Janke
University of Kent, UK

Marianne Gullberg
Lund University, Sweden

Chloë Ruth Marshall
University College London, UK

Abstract: Adults can extract phonological regularities from just several minutes' exposure to naturalistic input of an unknown spoken language (Gullberg et al., 2010). We examined whether such implicit statistical learning mechanisms also operate in the sign language modality. The input materials consisted of a continuous sign stream in the form of a weather forecast in Svenskt Teckenspråk (STS). L1-speakers of English with no prior knowledge of a sign language were assigned to two experimental groups who watched the forecast once (N=43) or twice (N=38), and a control group who did not watch it (N=40). Participants completed a 'surprise' lexical decision task designed to tap into their awareness of the phonological properties of the core STS lexicon. They viewed individual signs and indicated whether or not these could be real STS signs. The signs comprised four sets: STS signs that (1) were presented, and (2) were not presented, in the forecast; and signs that are not STS signs and (3) contain handshapes outside the STS handshape inventory, and (4) contain sets of phonological features that are dispreferred across sign languages. We found *no* evidence of any learning of STS phonological regularities. Considered in conjunction with two companion studies which *did* demonstrate some learning of sign forms and their meanings from these same input materials, our findings suggest limits to what can be learnt after just a few minutes of implicit and naturalistic exposure to language in an unfamiliar modality: information about specific lexical items is learnable, but information that requires generalisation across items may require greater amounts, or a different quality, of input.

Keywords: sign language; implicit learning; first exposure; phonological regularities; iconicity.

Corresponding author: Chloë Marshall, UCL Institute of Education, University College London, 20 Bedford Way, London WC 1H 0AL, UK. Email: chloe.marshall@ucl.ac.uk

ORCID ID: https://orcid.org/0000-0003-2405-1999

Citation: Hofweber, J., Aumônier, L., Janke, V., Gullberg, M., & Marshall, C. R. (2024). Can sign-naïve adults learn about the phonological regularities of an unfamiliar sign language from minimal exposure? *Language Development Research*, *4*(1), 455–480. http://doi.org/10.34842/hofweber

Introduction

The universality – or not – of language acquisition mechanisms has far-reaching implications for the disciplines of psycholinguistics, theoretical linguistics and language pedagogy. Although a substantial body of research has investigated language acquisition in both children and adults, the focus has been almost exclusively on spoken languages (Kidd & Garcia, 2022; Schönström, 2021). Sign languages – which are perceived and produced in the visuo-gestural modality – have been relatively neglected, meaning that little is known about the extent to which their acquisition and developmental trajectory resemble spoken languages. This is problematic because theories of first and second language acquisition based solely on spoken languages (and on the written form of some of those languages) make universal claims, yet it is not known whether those theories hold for sign languages too (Gullberg, 2022; Hou & Morford, 2020; Lillo-Martin & Hochgesang, 2022).

The focus of the current paper is on adult language acquisition mechanisms. An important issue in spoken second language acquisition research concerns how adults are able to learn new languages implicitly, and the type of linguistic knowledge they are able to acquire in this way. We extend this inquiry to sign language acquisition, and specifically to whether sign-naïve adults can learn about the phonological regularities of a target sign language at first exposure. In this introduction we review literature on implicit language learning at first exposure, before summarising the key findings of two studies on implicit sign language learning at first exposure that relate directly to the current one and so help situate and motivate it.

Implicit Language Learning on First Exposure to Naturalistic Input

In the field of adult second language acquisition, distinctions have traditionally been made between explicit and implicit learning. Explicit learning (often instructed learning in classrooms) is associated with intentional, deliberate attempts to memorise something with control, effort, and awareness, for example, in a setting where there has been advance warning of a test of learning. In contrast, implicit or incidental learning is characterised as learning without conscious attention to the input or effort, including in an experimental set-up in which participants are unaware that there will be a test of learning, and where they therefore have no intention to learn and no awareness of what is being learnt (e.g., Andringa & Rebuschat, 2015; DeKeyser, 2003; Godfroid, 2021; Hulstijn, 2005; Rebuschat & Williams, 2012; Williams, 2009, *inter alia*). The field is still rife with debate, especially concerning whether implicit learning can ever actually be tested since the very face of testing draws attention to language and potentially to learning. However, it is now widely recognised recognized that adult learning can take place under both explicit and implicit conditions. What is under discussion are the details, and which type of learning is optimal for which types of knowledge.

The current study focuses on *implicit* learning. We use the term simply to refer to learning taking place without instruction or training, recognising the difficulties of assuring that there is no conscious effort to learn or awareness of what is being learnt. Moreover, we focus on the very earliest stage of learning a new language, when the learner is a novice exposed to the language for the *first time*. Our interest is in how language learning gets off the ground, and specifically, how it does so when the language input is *naturalistic* in form. Second[1] language learning often takes place in instructed contexts, for example, in a classroom where linguistic input is broken down into manageable chunks, such as individual words, and often accompanied by explicit explanations, such as translations into the learner's first language. However, this is not the only way in which learners might encounter a new language. In many contexts, such as migration to another country for personal, political or economic reasons during adult life, learners might instead encounter a new language in more informal contexts. This will involve implicit learning through interactions with colleagues, watching television, listening to song lyrics, playing video games and engaging with social media (see pioneering studies by Meisel et al., 1981; Perdue, 1984; and studies in the emerging field of informal second language learning, e.g., Arndt, 2019; Dressman, 2020; Sockett, 2022). An interesting question is how learning takes place in these scenarios, where input is more continuous in nature, and what exactly can be learnt.

The process of breaking into a new language, what Klein (1986, p. 59) called the learner's "problem of analysis", contains three crucial aspects: (1) segmenting the continuous speech stream to identify relevant strings such as words, (2) identifying meaning that can be mapped onto those sound strings, and (3) generalising beyond the input exemplars to novel items so as to form linguistic categories and extract regularities. A set of three linked studies by Gullberg et al. (2010) investigated this very process by asking adult L1-speakers of Dutch to watch a 7-minute weather forecast presented in a typologically different, and previously unknown-to-them, language, namely Mandarin Chinese. Immediately after viewing the forecast, participants undertook 'surprise' tests tapping into form recognition, meaning assignment and phonotactic generalisations. Although participants found these tasks very challenging, and their overall performance was low, they did nevertheless show evidence of being able to extract word-form-related information, and of managing to extract lexical meaning from the context and map it onto word forms thus identified. They were also able to extract abstract, phonotactic information and generalize it to novel items not encountered in the input, a finding replicated by Ristin-Kaufman and Gullberg (2014) with the same materials but with Swiss-German speakers. These findings suggest that

[1] We use 'second language learning' to cover both second and foreign language learning, because the distinction between them does not matter for our purposes. Nor do we make a distinction in this article between 'acquisition' and 'learning'.

adult learners can deal efficiently and quickly with very complex language input at first exposure, even in the absence of instructions.

A question that arises, however, is whether such results would obtain in languages other than spoken Mandarin. For example, could they be found for a sign language in participants unfamiliar with sign languages? In the next section we review evidence from two recent studies that suggests they can.

Implicit Learning in a New Language Modality: Situating the Current Study

We adapted Gullberg et al.'s (2010) implicit learning paradigm to a sign language (specifically, Swedish Sign Language: Svenskt Teckenspråk, STS) to investigate whether the learning mechanisms identified in learners of spoken language at first exposure to uninstructed, naturalistic and continuous input are evident across the modality boundary. We created an STS weather forecast and produced STS versions of Gullberg et al.'s tasks for identifying word forms and lexical meaning. In two studies linked to the current study, we showed that (1) sign novices can distinguish between signs that they have and have not seen in the STS weather forecast, revealing that they are able to identify sign forms in the sign stream (Hofweber et al., 2022), and that (2) they are able to assign meaning to signs more accurately than control participants who have not viewed the forecast (Hofweber et al., 2023). As in Gullberg et al.'s studies, performance was not high[2], which indicates that the tasks are challenging. Nevertheless, the findings do provide evidence of implicit language learning.

The current study completes the trio of experiments by asking whether adults with no previous sign language exposure can generalise beyond encountered signed exemplars on first exposure. Can they, when presented with signs that they have not previously seen, make accurate judgements about which signs could be possible signs of the sign language they have viewed and those which could not? This question is particularly important with respect to implicit learning because whereas Hofweber et al.'s two previous studies focused on the encoding of lexical items in memory, here we ask whether sign novices can extract regularities across those items. Language acquisition – whether in an individual's first or subsequent languages – crucially involves learners being able to generalise beyond exemplars that they have encountered in the input, in order to form categories and establish regularities (see Ambridge, 2020, for an in-depth and up-to-date treatment of this topic). Hofweber et al. (2022, 2023) demonstrated that learners recognised and assigned meaning to new exemplars of lexical items viewed in the test phase of their studies. A remaining question, however, is whether they can establish regularities.

---

[2] In Study (1), accuracy rates ranged from 53% to 64% across conditions, with 50% representing chance performance, while in Study (2) participants were only able to generate the correct meaning for 14% of signs.

Recall that Gullberg et al. (2010) showed that L1-Dutch-speaking adults could extract phonotactic regularities, i.e., highly abstract information about sound structures, from input in another newly encountered spoken language (Mandarin Chinese). More specifically, learners rejected consonant-vowel-consonant forms with a phonotactically-illegal final consonant (e.g., *gam*). The ability to identify these words as impossible in Mandarin must have stemmed from participants analysing the new language input rather than from transferring their L1 phonotactic knowledge. This is because consonant-vowel-consonant words of this type are acceptable in Dutch (and, indeed, were accepted as possible Mandarin words by a control group of participants who had not viewed the weather forecast).

At this point, one might reasonably ask how 'phonotactics' - a term coined in spoken language linguistics for the rule-based ways in which phonemes can be combined - relates to sign languages. In fact, just as there are phonotactic constraints in spoken languages, where not all combinations of phonemes are possible, so are there constraints on how the formational units of signs (i.e., handshapes, movements and locations) are combined. Certain formational properties are dispreferred in lexical signs across the world's sign languages (Johnston & Schembri, 2009; Sandler, 2012; Sandler & Lillo-Martin, 2006). Examples of phonotactically-dispreferred signs often involve a change of handshape, i.e., the second handshape involves a different set of selected fingers to the first (thereby violating the 'selected fingers constraint', Mandel, 1981), or two moving hands have different handshapes (thereby violating the 'symmetry condition', Battison, 1978).

Preliminary evidence indicates that people might indeed be sensitive to phonotactic regularities in sign when exposed to sign language for the first time. Hofweber et al. (2022) conducted a recognition test, whereby participants first watched a 4-minute weather forecast in STS and were then presented with individual signs. Their task was to decide whether or not each of those signs had or had not appeared in the forecast. There were two sets of items that participants had not seen in the forecast. One comprised phonologically plausible signs in that they were real signs of STS that shared some phonological features with the target signs. The second set, however, were phonologically implausible in that although they were real signs in other sign languages, they were not STS signs and contained phonological features that are dispreferred across the world's sign languages (because they violated the selected fingers constraint or symmetry condition explained above). Participants were more accurate than chance at responding 'no' to both sets of signs that they had not seen in the forecast, but the fact that this effect was greater for the implausible signs than the plausible signs suggests some sensitivity to sign phonotactics.

Further with respect to sign language phonology and the phonotactic constraints discussed above, the core lexicon[3] of each sign language uses a specific set of handshapes. The handshape inventory of any particular sign language is smaller than the larger set of handshapes attested among sign languages of the world (just as each particular spoken language uses only a portion of the phonemes attested worldwide) (Brentari & Eccarius, 2010). This inventory may be augmented by handshapes that represent letters in the one-handed manual alphabet ('fingerspelling'). The inventory for STS can be found here https://teckenspraakslexikon.su.se/handformer. We investigated whether our participants would be sensitive to the handshapes that had occurred in the weather forecast so, when shown signs that they had *not* seen in the input materials, they would accept signs containing handshapes they had seen as possible signs of the language but reject signs containing handshapes they had not seen. Because in the current study we look at both phonotactics and handshapes, we use the broader term 'phonological regularities' from here on.

For the current task, we used the same STS weather forecast exposure video as in Hofweber et al. (2022, 2023) but administered a different experimental task to determine whether sign novices are sensitive to phonological regularities in a newly encountered sign language. Our experimental task required participants to decide whether the signs being shown were real signs of STS or not. There were four sets of signs. Two of these sets comprised real STS signs, one of which appeared in the forecast and one of which did not. The remaining two sets comprised signs that are not signs of STS (and therefore which also did not appear in the forecast). More specifically, one of these sets comprised signs with handshapes outside the STS handshape inventory, and the other comprised signs with sets of phonological features that are dispreferred across sign languages (because they violated the selected fingers constraint or symmetry condition).

Of course, it might be the case that people who have never experienced a sign language nevertheless have expectations of what signs might look like. After all, they bring with them a lifetime of using their hands to make co-speech gestures and of watching the co-speech gestures of others. Furthermore, there is growing evidence that the phonological system is, at least in part, amodal, and that sign-naïve adults can transfer knowledge of regularities in their spoken L1 to signs (Berent & Gervain, 2023, and references therein). It is possible that certain hand configurations and movements are more plausible than others when people are required to make an explicit

[3] The lexicon of many sign languages, including STS, has a tripartite lexicon, comprising core signs, non-core (depicting and pointing) signs and borrowed (fingerspelt) signs (Johnston & Schembri, 2009). Phonological constraints are most strongly attested in the core lexicon (Brentari & Eccarius, 2010). Apart from the occasional pointing sign (pronouns signed with the index finger, and flat-handed points to the weather map), our weather forecast contains just core lexical signs.

judgement on what could or could not be a sign. To control for this possibility, a group of participants undertook the lexical decision task without having watched the weather forecast.

We expected the task to be challenging (as it was for the participants in the studies by Gullberg et al., 2010; Hofweber et al., 2022; Hofweber et al., 2023; Ristin-Kaufmann & Gullberg, 2014) but we predicted that participants who had viewed the forecast would be more accurate at distinguishing between STS and non-STS signs than those who had not viewed it. Importantly, if participants were able to accept both real STS signs that they had just viewed and signs that they had not, and also reject signs that are not signs of STS (which, by definition, would not have appeared in the forecast), this would indicate some level of generalization across the input and not just the recognition of viewed exemplars.

Methods

The materials used in this study can be accessed on the Open Science Framework site via the following link: https://osf.io/8hrp6/. It should be noted that video materials need to be downloaded for viewing.

Participants

This study was originally designed to be run face-to-face in the lab. However, due to testing restrictions in place during the Covid-19 pandemic, we adapted it for online administration. Hence, we uploaded the experiment originally designed in PsychoPy onto the Pavlovia platform, https://pavlovia.org/. This allowed us to reach a more diverse participant pool than is common in lab-based psychology studies. However, the online format also meant that compatibility issues between Pavlovia and participants' domestic software set-up meant that data from some participants (7 in the 0x, 10 in the 1x, and 3 in the 2x Exposure groups) could not be collected. Hence, this paper only reports data collected from participants in which no compatibility issues occurred and the full experimental task could be administered (final *N*=121). In addition, some of the Wechsler Adult Intelligence Scale (WAIS) vocabulary task data had to be discarded (data from 3 participants each in the 0x and 2x Exposure groups, and 15 participants in the 1x Exposure group) due to problems with the audio quality of the voice recording that made transcription unreliable.

Participants were recruited using the website "Call for participants" (https://www.callforparticipants.com). None of them had participated in either of the companion studies to this study (Hofweber et al., 2022; Hofweber et al., 2023). They were adult native-speakers of English with no prior knowledge of any sign languages or of Swedish. Given that the tasks were visual, and to avoid confounds from age-related decline of vision, the maximum age for participants was set at 40 years.

Participants were randomly assigned to one of the following three groups, with group sizes designed to be twice as large as those used by Gullberg et al. (2010) and commensurate with those used by Hofweber et al. (2022) and Hofweber et al. (2023):

- 1x exposure group (*N*=43; 17 males): exposed to the STS forecast once
- 2x exposure group (*N*=38; 19 males): exposed to the STS forecast twice
- 0x exposure group (*N*=40; 21 males): control group who were not exposed to the STS forecast.

Using a detailed demographic and language background questionnaire, each participant was assessed for general demographic variables, such as age and education (an indicator of socio-economic status), as well as for existing knowledge of languages. We also assessed non-verbal reasoning using the matrices subtest of the WAIS III, and English (i.e., the participants' L1) vocabulary knowledge using the vocabulary subtest of the WAIS IV. Between-subject ANOVAs were conducted to compare the participant groups on these background variables. As can be seen in Table 1, the three groups were matched for age, education and number of known languages. However, the 1x exposure group displayed significantly higher non-verbal reasoning and English vocabulary scores than the other groups. Nevertheless, these factors did not correlate with the dependent variable of the study (accuracy in the experimental task), so the difference was not interpreted further.

Table 1. ***Participant background variables***

		Exposure group			*F*	*p*	η^2
		0x	1x	2x			
Age (years)	Mean	23.87	25.88	24.30	2.06	.13	.05
	SD	2.56	5.10	3.51			
Education (years)	Mean	16.53	16.58	16.97	0.23	.79	.005
	SD	2.79	3.37	2.21			
English vocabulary knowledge (WAIS IV vocabulary raw score)	Mean	33.17	38.42	34.94	3.93	.02*	.09
	SD	6.85	7.31	6.63			
Number of known languages	Mean	1.87	2.04	2.24	0.58	.57	.01
	SD	1.38	1.16	1.54			
Non-verbal reasoning ability (WAIS III matrices raw score)	Mean	18.67	22.21	19.73	5.83	.004**	.12
	SD	3.48	4.38	3.75			

Key: * = p < .05; ** = p < .01

Experimental Materials

Weather Forecast

The exposure materials consisted of a 4-minute weather forecast in STS, recorded by a hearing native signer of STS who is a qualified and highly experienced interpreter. These materials were created specifically for this study and for those reported in Hofweber et al. (2022, 2023). Because our participants had English as their L1 and because we were particularly interested in the manual aspects of sign language, we did not use British Sign Language (BSL): the mouthing of many BSL signs is based on the lip patterns for the corresponding English words, so our participants might have relied on lip-reading to make sense of the input materials. For this reason, we chose another language – STS – as the target language.

Amongst a range of other signs, the forecast incorporated 22 target signs which featured in the lexical decision task (designed to be similar in number to the 24 target words used by Gullberg et al., 2010, in their Mandarin weather forecast). These 22 target signs occurred with different frequencies in the forecast. We presented 11 'high frequency' target signs (8 occurrences) versus 11 'low frequency' target signs (3 occurrences; with the exception of one item, SÖDER 'south', which occurred 4 times in error). The high and low frequency sets were matched for a range of crucial aspects of sign language, such as phonology (i.e., locations and hand configurations, and the number of one-handed versus two-handed signs) and iconicity. Iconicity ratings collected from an independent group of 24 sign-naïve participants confirmed the experimenters' intuition that high (M = 3.64, SD = 1.55) and low frequency (M = 3.68, SD = 1.76) signs did not differ in their level of perceived iconicity, $F(1,22) = 0.003$, $p = .96$, $\eta^2 = 0.000$ (see Hofweber et al., 2023, for further details). The two types of target items were also matched for the occurrence frequency of their English translation equivalents using CELEX corpus (Baayen et al., 1995): low: M = 32,759, SD = 51,978; high: M = 27,027, SD = 22,771, $F(1,22) = 0.11$, $p = .74$, $\eta^2 = 0.006$.

Lexical Decision Task

Participants saw 88 short video clips of individual signs, each signed by the same signer who signed the forecast. After each clip, they made a meta-linguistic judgement as to whether or not the sign was a real sign of sign languages. All participants were presented with the same video stimuli but the instructions differed slightly, depending on whether participants had seen the forecast or not. Participants in the 1x and 2x exposure groups were asked if the signs could be real signs of STS specifically, whilst participants in the control group were asked whether the signs could be real signs of sign languages more generally. Control participants received different instructions because it would have been pragmatically odd to ask them to make a judgement relating to a specific sign language they had never seen before, but they could be assumed to have some expectation of what sign languages in general look like.

The stimuli consisted of four sets of signs. None of these signs contained mouthings (i.e., silent mouth patterns from spoken words that signers sometimes use to accompany manual signs), so participants were not able to gain any (spoken) language information from viewing the signer's mouth. Set 1 included 22 stimuli that were the target signs from the forecast, so the accurate answer to these was 'yes', given that they are real signs of STS. Set 2 had 22 stimuli which were also real signs of STS. They contained handshapes, hand orientations, movements and locations that had occurred in the forecast but that were combined in different ways from those shown in the forecast. Therefore, for this set, too, the correct response was 'yes'. The remaining two sets of signs required the response 'no'. Set 3 comprised 22 signs with handshapes (different for each sign) that are not part of the core STS lexicon, and had therefore not appeared in the forecast. However, these handshapes do occur in other sign languages and the signs were indeed real signs (e.g., American Sign Language, Chinese Sign Language, Kenyan Sign Language and Khmer Sign Language). The fourth and final set comprised 22 signs with the sorts of handshape changes and movements that violate constraints of sign formation and are therefore dispreferred across the world's sign languages. None of the signs shown in the forecast violated these constraints. Again, however, signs in this set were real signs from other sign languages. Table 2 summarises the properties of each set of signs. Note that all other phonological differences between the four sets were minimal: all four sets were matched for the number of one- and two-handed signs, and Set 2 was matched to Set 1 for the hand configurations used.

Supplementary Tasks

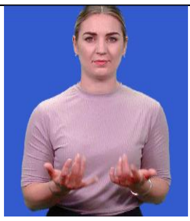
English Vocabulary Knowledge

Participants' knowledge of their first language was measured using the English vocabulary subtest of the Wechsler Adult Intelligence Scale WAIS-IV (Wechsler et al., 2008; not available on the osf site because it is proprietary). In this test, participants are presented with 26 lexical items auditorily and orthographically and asked to define each one. Their responses were recorded using the audio software *Audacity* and scored based on the test manual. The responses from a subset of participants (N=10) were scored by two independent judges, resulting in an interrater reliability score of Spearman's Rho = .85, p = .002.

Non-verbal Reasoning Ability

The matrices subtest of the Wechsler Adult Intelligence Scale WAIS-III (Wechsler, 1997; not available on the osf site because it is proprietary) taps into individuals' visual ability to recognise patterns. Participants view designs of shapes and colours. Each design contains a gap. In a multiple-choice style, participants choose one from several options to complete the design. The task was administered online using google docs.

Table 2. ***Properties of the stimulus items in each sign set***

Set	Signs are included in the weather forecast	Signs are in the core STS lexicon	Signs use handshapes from the core STS lexicon	Signs violate the phonotactic constraints of sign languages	Example (photos illustrate each sign's hand configuration(s) and location, and the movement is described in the text. The item ID can be used to locate the video on the osf site)
1	yes	yes	yes	no	 Sweeping movement across chin Gloss: This sign means 'warm' in STS THF1
2	no	yes	yes	no	 Repeated finger wiggle Gloss: This sign means 'simmer' in STS LDD1THF3
3	no	no	no	no	 Double short downward movement Gloss: This sign means 'Namibia' in Namibian Sign Language. LDD2THF1
4	no	no	yes	yes	 Repeated asynchronous movement of the two hands that have different handshapes Gloss: This sign means 'SimCom' in American Sign Language LDD3THF2

Procedure

Since the data collection for this study took place in 2021 when in-lab testing was not permitted due to the Covid-19 pandemic, the study was conducted online using the Microsoft Teams software. The experimenter met and observed each participant individually on Teams. Participants in the exposure groups first watched a short video of a weather forecast in STS. Because the study was designed to tap implicit learning, the instructions were minimal to avoid explicit reference to learning. Participants were simply told to watch the signer as she signed the forecast. Immediately after watching the forecast, they completed the surprise lexical decision task. The control group proceeded with the lexical decision task straight away, without having watched the forecast first. Upon completion of the main experimental tasks, participants completed the WAIS non-verbal reasoning and vocabulary tests. Finally, they filled in the demographic and language background questionnaire on Surveymonkey.

Data Analysis

To analyse the lexical decision task results, we investigated both accuracy rates and yes responses for each item and participant. The summary tables of our results are presented using the style adopted by Ortega et al. (2019). The full data set is available on the osf site: https://osf.io/8hrp6/. Generalised mixed model analyses were conducted in R studio using the lmer.test package in R (Kuznetsova et al., 2017), which automatically generates significance levels for each effect. We initially assumed a maximally conservative approach to random effects, allowing both items and subjects to vary by both intercept and slope (Winter, 2019). However, this resulted in failure to converge due to the model complexity, so we simplified the models to vary only by intercept. The alpha level was set at .05. Due to limitations regarding modelling pair-wise comparisons involving variables with three or more levels in lme4 (Winter, 2019), we conducted the analyses in several steps.

The first step was to assess the overall effect of exposure by comparing results in the control group (0x) to results in the two exposure groups (1x, 2x). Secondly, we compared the two exposure groups to each other. Finally, we assessed the effect of Set for exposure and non-exposure groups separately. Since, there were no differences in pattern across the 1x and 2x exposure groups, the two groups were combined for the analyses by Set. In these analyses by Set, the intercept was set as the values of Set 1 (i.e., target) items.

The analyses on effects of Set were conducted separately for Accuracy and Yes responses, to reveal response biases. Any differences in accuracy between the two sets of items that required a ‘yes’ response (Sets 1 and 2) and the two sets that required a ‘no’ response (Sets 3 and 4) could potentially be driven by participants not actually making a distinction between any of the sets at all, and responding ‘yes’ at similarly high levels throughout, regardless of the phonological properties of the signs. In order to investigate this possibility, we repeated the analysis using Yes responses rather than Accuracy as the dependent variable. We note that the variable 'Accuracy' makes less sense for the control group who did not see the forecast, because their task instructions asked not about STS specifically but about sign languages more generally; because the signs used in Sets 3 and 4 were real signs (albeit ones that are formationally rare) the correct response for these signs was - like it was for Sets 1 and 2 - 'yes'. However, we include the data from the control group in the Accuracy analyses for the sake of completeness.

Results

Results for Accuracy Rates

Table 3 presents the descriptive statistics for Accuracy rates by exposure group and sign set. Figures 1 and 2 illustrate the distribution of accuracy rates. Tables 4, 5, 6 and 7 present the inferential statistics based on linear mixed effects models. Table 4 presents the effect of overall exposure on accuracy. Table 5 presents the effect of number exposures (1x versus 2x) on accuracy. Table 6 presents the effects of Set on accuracy in the 0x control group. Table 7 presents the effects of Set on accuracy in the two exposure groups. Based on Tables 4 and 5, accuracy did not differ by exposure group. As a result, subsequent analyses did not differentiate between 1x and 2x exposure groups. However, the different sign sets yielded different accuracy rates. Whilst there were no accuracy differences between Sets 2 and Sets 1 (target signs), accuracy in Sets 3 and 4s was lower than in Set 1 (target signs). This effect applied across all exposure groups.

Table 3. ***Accuracy rates by exposure group and sign set***

		0x Exposure group	1x Exposure group	2x Exposure group
Accuracy rate (%)	Mean	59.52	55.60	61.41
Set 1 signs	SD	23.07	16.73	16.78
Accuracy rate (%)	Mean	59.57	61.91	59.50
Set 2 signs	SD	24.75	13.86	17.11
Accuracy rate (%)	Mean	42.74	44.64	46.87
Set 3 signs	SD	22.40	16.20	19.81
Accuracy rate (%)	Mean	40.63	44.55	47.12
Set 4 signs	SD	19.63	16.15	19.87

As explained in the text, accurate responses for Sets 1 and 2 were 'yes' and for Sets 3 and 4 were 'no'.

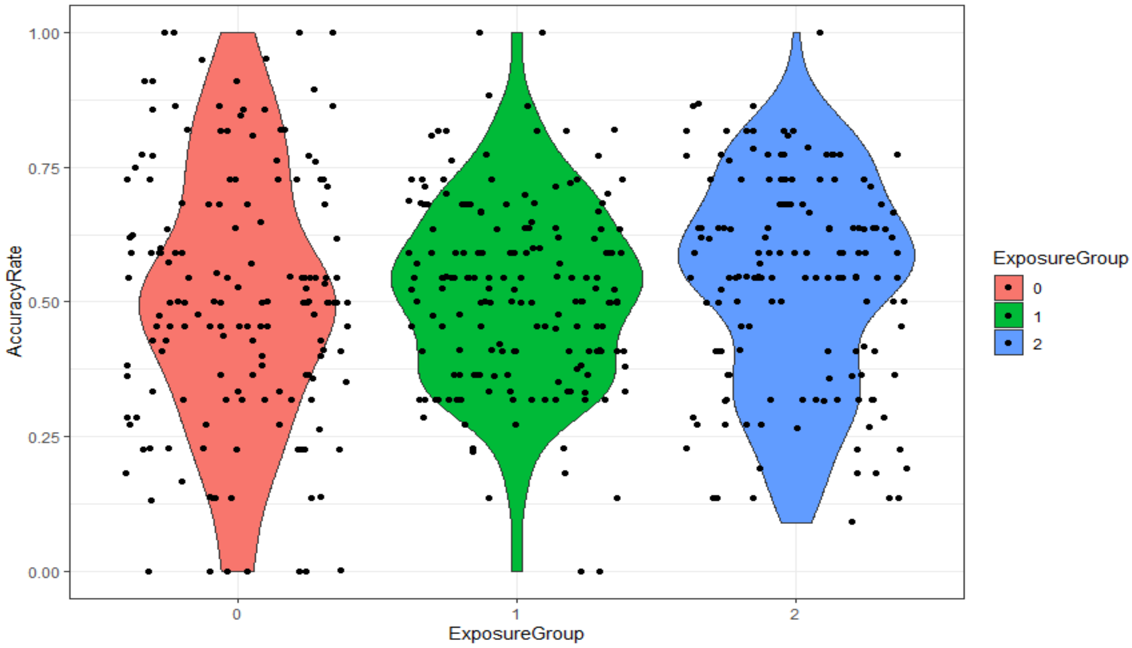


Figure 1. ***Accuracy rates by exposure group***

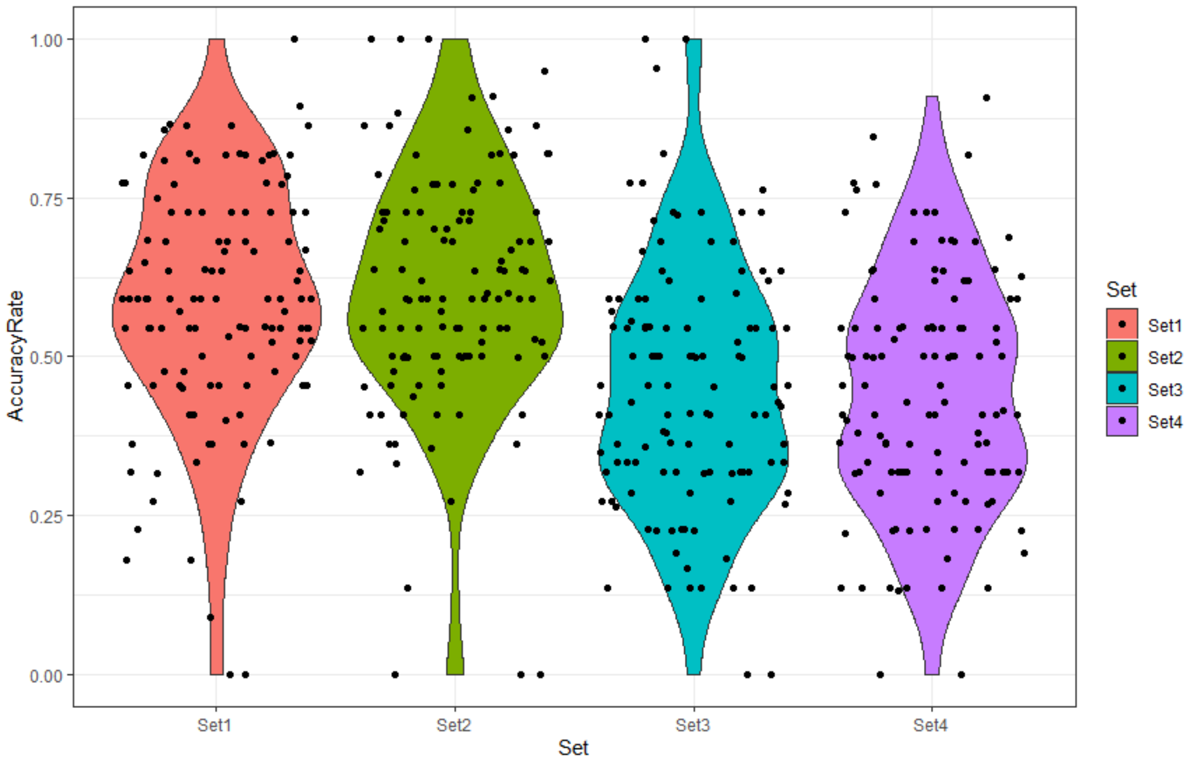


Figure 2. ***Accuracy rates by set***

Table 4. ***Model output for Accuracy: Effect of Exposure (0x versus 1x/2x)***

Predictors	β	SE	Z	*p*
Intercept (0x group)	0.04	0.07	0.67	.50
Exposure (1x/2x groups)	0.01	0.03	0.33	.74

glmer (accuracy ~ exposure+(1|item)+(1|subject),data=Data, family=binomial)

Table 5. ***Model output for Accuracy: Effect of Exposure times (1x versus 2x)***

Predictors	β	SE	Z	*p*
Intercept (1x group)	0.01	0.08	0.17	.86
2x exposure	0.09	0.08	1.13	.26

glmer (accuracy ~ exposure_times+(1|item)+(1|subject),data=Data_exp, family=binomial)

Table 6. ***Model output for Accuracy: Effect of Set in 0x exposure group***

Predictors	β	SE	Z	*p*
Intercept (set 1)	0.44	0.08	5.52	<.01
Set 2	0.05	0.11	0.43	.67
Set 3	-0.80	0.11	-7.61	<.01
Set 4	-0.86	0.11	-8.12	<.01

glmer (accuracy ~set+(1|item)+(1|subject),data=Data_0, family=binomial)

Table 7. ***Model output for Accuracy: Effect of Set in 1x/2x exposure group***

Predictors	β	SE	Z	*p*
Intercept (Set 1)	0.32	0.10	3.04	<.01
Set 2	0.15	0.14	1.08	.28
Set 3	-0.51	0.14	-3.66	<.01
Set 4	-0.67	0.14	-4.79	<.01

glmer (accuracy ~set+(1|item)+(1|subject),data=Data_exp, family=binomial)

Additional analyses were conducted to explore the predictors of Accuracy in Set 1 items. No significant effects of input-related factors, such as frequency or iconicity (p-values > 0.05), or correlations with individual differences (age, education, non-verbal reasoning ability, English vocabulary, number of languages known) were observed (all r-values < .10).

Results for Yes responses

Table 8 presents the descriptive statistics for Yes response rates by exposure group and set. Figures 3 and 4 illustrate the distribution of Yes responses. Whilst Tables 9 and 10 present the inferential statistics based on linear mixed effects models. the effect of group will be identical for accuracy and Yes response rates, the effect of Set may not be. Table 9 presents the effects of Set on Yes response rates in the 0x control group. Table 10 presents the effects of Set on Yes response rates in the two exposure groups. In neither analysis was there a significant effect of Set, indicating that participants – whether or not they had watched the forecast – did not respond differently to signs that were or were not signs of STS. There is therefore no evidence for the learning of phonological regularities during the viewing of the input materials.

Table 8. ***Yes response rate by exposure group and sign set***

		0x Exposure	1x Exposure	2x Exposure
Yes rate (%)	Mean	59.52	55.60	61.41
Set 1 signs	SD	23.07	16.73	16.78
Yes rate (%)	Mean	59.57	61.91	59.50
Set 2 signs	SD	24.75	13.86	17.11
Yes rate (%)	Mean	57.26	55.36	53.13
Set 3 signs	SD	22.40	16.20	19.81
Yes rate (%)	Mean	58.37	59.84	56.88
Set 4 signs	SD	23.54	16.93	18.46

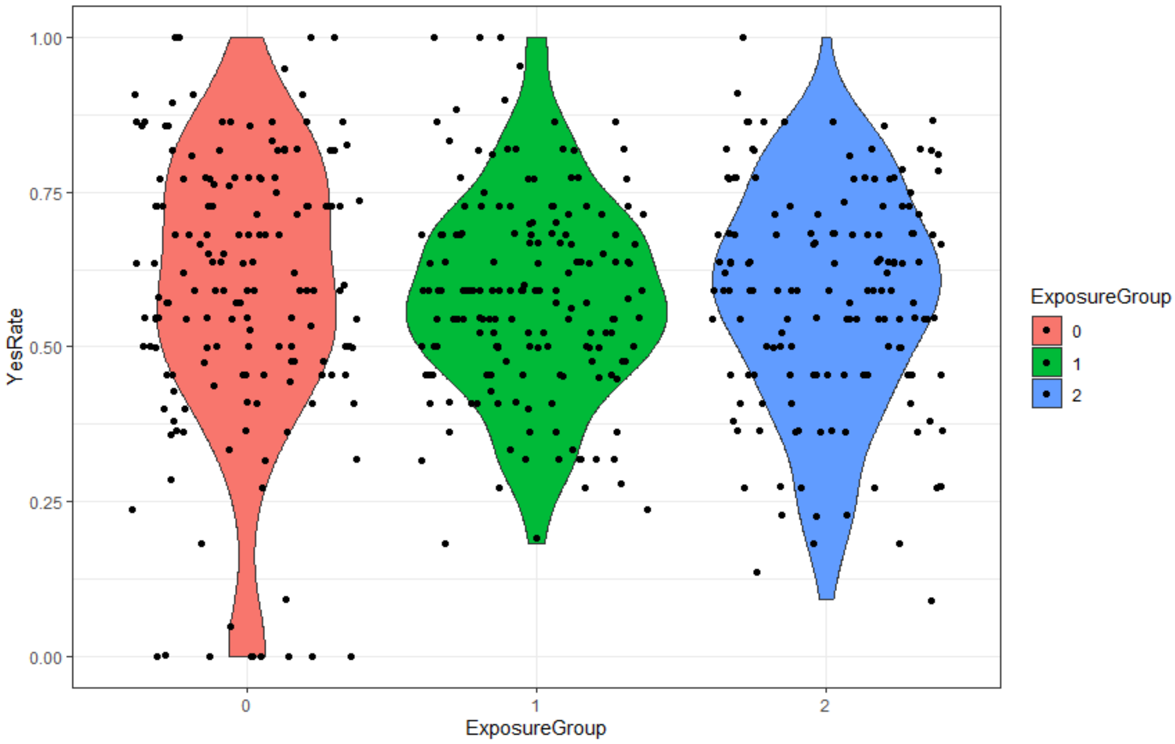


Figure 3. ***Yes rates by exposure group***

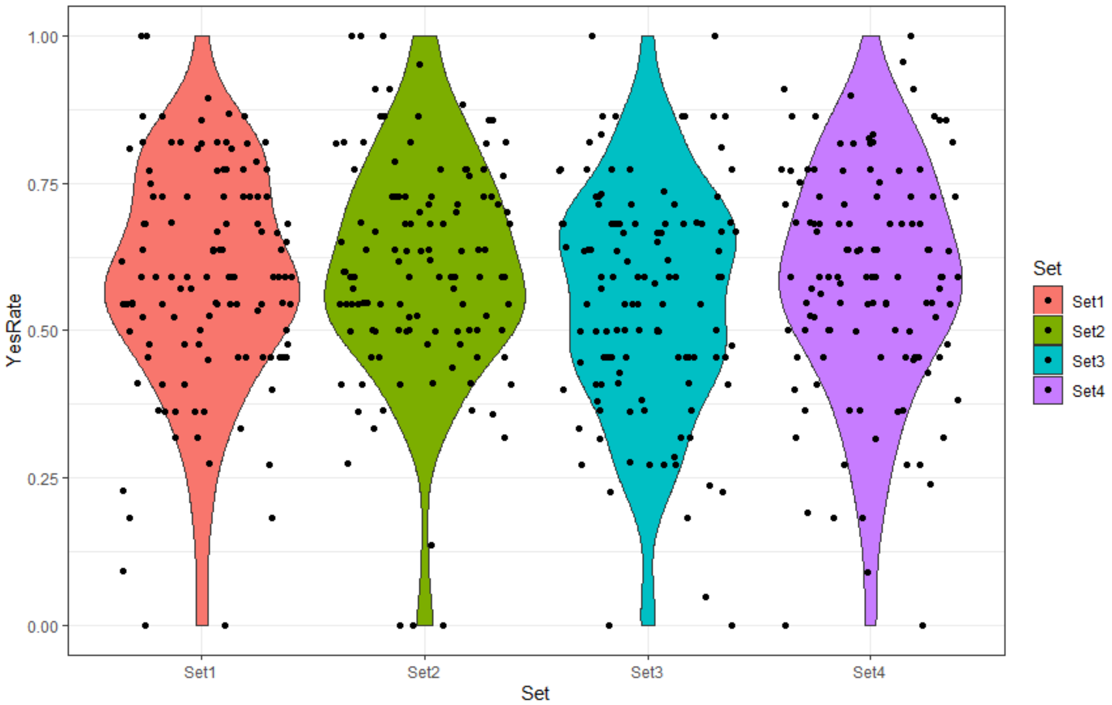


Figure 4. ***Yes rates by set***

Table 9. ***Model output for yes response rate: Effect of set in 0x exposure group***

Predictors	β	SE	Z	*p*
Intercept (Set 1)	0.40	0.20	2.02	.04
Set 2	0.03	0.12	0.23	.82
Set 3	0.09	0.12	-0.73	.47
Set 4	-0.03	0.12	-0.25	.81

glmer (response1 ~set+(1|item)+(1|subject),data=Data_0, REML=false)

Table 10. ***Model output for yes response rate: Effect of set in 1x/2x exposure group***

Predictors	β	SE	Z	*p*
Intercept (Set 1)	0.34	0.12	2.85	<.01
Set 2	0.16	0.15	1.07	.29
Set 3	-0.13	0.15	-0.86	.38
Set 4	0.04	0.15	0.30	.76

glmer (response1 ~set+(1|item)+(1|subject),data=Data_exp, REML=false)

Discussion

In this study, we investigated whether sign-naïve adults are able to learn about the phonological regularities of a target sign language at first exposure in an implicit, uninstructed learning context. The input materials consisted of a continuous sign stream in the form of a 4-minute weather forecast video in Swedish Sign Language (STS). After having watched the forecast either once or twice, participants completed a 'surprise' lexical decision task designed to tap into their understanding of the phonological properties of STS and sign languages in general. The participants viewed individual signs and were asked to indicate whether each sign could be a 'real sign of STS'. Stimulus items comprised four sets: (1) STS signs that were presented in the forecast, (2) STS signs that were not presented in the forecast, (3) signs that are not STS signs and contain handshapes outside the STS handshape inventory, (4) signs that are not STS signs and contained phonological features that are dispreferred across the world's sign languages. Correct answers to Sets 1 and 2 were 'yes' and to Sets 3 and 4 were 'no'. We also tested a group on the experimental task who had not viewed the forecast.

Our predictions were not borne out by the data. First of all, we had predicted that being exposed to an unfamiliar language in an unfamiliar modality – albeit under implicit learning conditions and for only a short time – would lead to greater accuracy in distinguishing between signs that were STS signs versus those that were not, in comparison to participants who had had no exposure. In other words, we had predicted that participants would be able to learn something about the phonological regularities of sign in a short period. However, this was shown not to be the case. In all three groups (0x exposure, 1x exposure, 2x exposure), the rate of accepting signs as possible signs of STS was around 58%. In fact, this rate of acceptance was fairly consistent across different sign sets too (Set 1, 59%; Set 2, 60%; Set 3, 55%; Set 4, 58%). In other words, participants did not distinguish between the different sets of signs in their responses, and so did not behave as we had expected – they did not reject the two sets of signs that we had predicted would be phonologically implausible to them. We had hypothesized that if participants who had viewed the weather forecast were able to accept both real STS signs that they had just viewed in the forecast and signs that they had not, and also reject signs that are not signs of STS (which, by definition, would not have appeared in the forecast), then this would indicate some level of generalization across the input and not just the recognition of viewed exemplars. We did not find evidence to support this hypothesis.

Taking the results of this study together with our two earlier studies (Hofweber et al., 2022; Hofweber et al., 2023), we appear to have found some limits on what can be learnt of a sign language at first exposure to brief and naturalistic input: lexical information – namely sign forms, and the meaning of signs – can be learnt, but it appears that phonological regularities cannot. This is in interesting contrast to Gullberg et al.'s spoken language studies of Mandarin Chinese learning by Dutch and Swiss-German speakers (Gullberg et al., 2010; Gullberg et al., 2012; Ristin-Kaufmann & Gullberg, 2014), where phonotactic restrictions on syllable-final consonants *were* learnt. And yet, it is not the case that our participants were responding 'yes' and 'no' at chance: we found a bias towards responding 'yes'. Participants were therefore erring on the side of being more rather than less accepting of the types of signs that could be part of the STS lexicon. It appears that just four minutes of exposure to naturalistic input is below the threshold for any learning of phonological regularities to occur.

And yet, we did have some hints from a previous study that learners might be able to extract phonological regularities after brief exposure to the same signed input materials. In Hofweber et al.'s (2022) recognition study – where participants had to make a decision as to whether signs had or had not appeared in the weather forecast – there were two sets of items that they had not seen in the forecast. One set comprised signs that were phonologically plausible in that they were real signs of STS and they shared some phonological features with the target signs. The second set of signs, however, were phonologically implausible because they contained phonological features that are dispreferred across the world's sign languages (like the signs in Set 4 in the current study, they broke the selected fingers constraint or symmetry condition). Hofweber et al.'s (2022) participants were more accurate than chance at responding 'no' to both sets of signs that they had not seen in the forecast, but this effect was greater for the implausible signs than the plausible signs, suggesting some sensitivity to phonological regularities.

The difference in findings could perhaps be explained by differences in task and instructions. The task in the current study required higher levels of meta-linguistic awareness. Whilst in Hofweber et al.'s (2022) study, participants were simply asked whether they recognized the sign from the forecast, in the current study they were asked to make a complex and abstract judgement, i.e., 'could this be a real sign of Swedish Sign Language or not'. It is possible that this type of judgement is too challenging for someone with no expertise in language studies. Moreover, it may have been difficult for participants without any prior experience with sign languages to make a particular judgement relating to one sign language. This would have been less of a consideration for speakers being asked about a particular spoken language such as Mandarin Chinese, as was the case in Gullberg and colleagues' studies. Their speakers already had familiarity with spoken languages and will have developed an awareness of the fact that languages can 'sound different' from each other. Can we assume the same for someone who has never learnt a sign language? After all, they were being exposed not just to an unfamiliar language but to an unfamiliar modality. Maybe they interpreted 'could this be a real sign of Swedish Sign Language or not' as a question about modality rather than a particular language, e.g., 'could this be a real sign of a sign language or not' (i.e. the question that the control group, who did not view the forecast, were asked), in which case it might seem strategic to be relatively generous with what can be accepted. That might also explain why the exposure groups did not differ from the control group in their responses.

An alternative explanation of our findings is that the phonological constraints posited for the core lexicon of sign languages are not as strong as is assumed in the literature. It is certainly true that these constraints are broken outside the core lexicon in classifier constructions and in signs that incorporate elements of fingerspelling (Johnston & Schembri, 2009; Sandler & Lillo-Martin, 2006). Furthermore, there is a greater range of handshapes outside the core lexicon (Brentari & Eccarius, 2010). Therefore, it might be that learners need longer than just a few minutes of exposure to learn how these constraints apply to lexical signs. A limitation of our study is that we did not check how strong these phonological constraints are for native signers of STS: we did not investigate how sign-like (or un-sign-like) they would judge the signs in Sets 3 and 4 to be, and this would be a useful addition to any future studies using our experimental paradigm.

A further consideration is whether participants paid sufficient attention to the weather forecast and the experimental task. If not, that could have contributed to our lack of differences between the different sign sets. However, two points suggest that poor attention is not necessarily an issue: (1) in the online presentation of our companion study tapping lexical meaning (Hofweber et al., 2023), a learning effect *was* obtained, and (2) the 'yes bias' in the current study suggests that participants were not responding completely at chance and were trying to respond accurately. With hindsight, the inclusion of catch trials would have been useful in order to determine more directly whether our participants were paying attention, and to exclude any who were not.

An interesting next step would be to explore whether, in contrast to non-signers, people with experience of a sign language unrelated to STS are able to learn its phonological regularities from our same input materials. As Chen Pichler and Koulidobrova (2023) discuss, the current literature on sign language learning has focused on hearing adults learning their first language, but deaf signing adults who are learning a new sign language are a key group for fully understanding the impact of modality on second language learning. In the case of the present task and those reported in our related papers (Hofweber et al., 2022; Hofweber et al., 2023), such a group would allow us to disentangle second-language learning effects within the signed modality from learning effects in a new modality.

The findings from our study also raise questions about possible differences between child and adult acquisition mechanisms, and between first and second language acquisition of phonotactic regularities in a visual modality. In the domain of spoken language, a large body of work has shown that both children and adults are able to successfully extract phonotactic regularities from spoken/auditory input (see Frost et al., 2019, for an overview), although most studies have operated with training paradigms rather than exposure to continuous input without training (but see Ristin-Kaufmann & Gullberg, 2014). Much less is known about differences between children and adults for the implicit learning of sign language phonotactics, let alone in a statistical learning paradigm. We know of no studies that test children's capacity for extracting phonotactic regularities in the visual modality, which means that we cannot tell whether the adults in this study are worse at this task than children. It therefore remains an empirical question with interesting theoretical ramifications to compare first and second language learners in this domain.

In conclusion, we found no evidence that hearing adults, after brief exposure to an unfamiliar language in an unfamiliar modality, were able to demonstrate learning of the phonological regularities explored in our study. Considered in conjunction with two companion studies revealing that participants *were* able to demonstrate learning of sign forms and their meanings after viewing these same input materials, we argue that our findings demonstrate the limits of what can be learnt: information about specific lexical items is learnable, but information that requires generalisation across items may require greater quantities of input or a different quality of input. All three of our studies need replication, preferably with different input materials, to establish their robustness. Furthermore, different conditions that might support the learning of phonological regularities need to be explored, for example, longer exposure time and explicit pointers or explanations, and learning might be better demonstrated using different tasks. Finally, we acknowledge that evidence for the implicit learning of phonological regularities at first exposure to an unfamiliar spoken language rests on just two studies of Chinese, and that these findings need to be replicated in other spoken languages (and preferably in languages with phonotactic properties that are very different to those of Chinese). Taken together, such studies in signed and spoken languages will help clarify the extent to which adult language acquisition mechanisms operate similarly or differently across modalities.

References

Ambridge, B. (2020). Against stored abstractions: a radical exemplar model of language acquisition. *First Language, 40,* 509-559. https://doi.org/10.1177/0142723719869731

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning. *Studies in Second Language Acquisition, 37,* 185-96. https://doi.org/10.1017/S027226311500008X

Arndt, H.L. (2019). *Informal second language learning: The role of engagement, proficiency, attitudes, and motivation* [PhD thesis]. University of Oxford.

Battison, R. (1978). *Lexical borrowing in American Sign Language*. Linstok Press.

Berent, I., & Gervain, J. (2023). Speakers aren't blank slates (with respect to phonology)! *Cognition, 232,* 105347. https://doi.org/10.1016/j.cognition.2022.105347

Brentari, D., & Eccarius, P. (2010). Handshape contrasts in sign language phonology. In Brentari, D. (Ed.) *Sign languages* (pp. 284-311). Cambridge University Press.

Chen Pichler, D., & Koulidobrova, E. (2023). The role of modality in L2 learning: The importance of learners acquiring a second sign language (M2L2 and M1L2 learners). *Language Learning, 73,* Issue S1, 197-233. https://doi.org/10.1111/lang.12607

DeKeyser, R.M. (2003). Implicit and explicit learning. In C.J. Doughty & M.H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Blackwell.

Dressman, M. (2020). Introduction. In M. Dressman & R.W. Sadler (Eds.) *The handbook of informal language learning* (pp.1-12). Wiley.

Ellis, N.C. (2012). What can we count in language, and what counts in language acquisition, cognition, and use? In S. T. Gries & D. Divjak (Eds.), *Frequency effects in language learning and processing* (Vol. 1, pp. 7-34). Mouton de Gruyter.

Frost, R., Armstrong, B.C., & Christiansen, M.H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin, 145*(12), 1128-1153. https://doi.org/10.1037/bul0000210

Godfroid, A. (2021). Implicit and explicit learning and knowledge. In H. Mohebbi & C. Coombe (Eds.), *Research questions in language education and applied linguistics* (pp. 823-829). Springer.

Gullberg, M. (2022). Why the SLA of sign language matters to general SLA research. *Language, Interaction and Acquisition, 13*(2), 231-253. https://doi.org/10.1075/lia.22022.gul

Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., & Indefrey, P. (2010). Adult language learning after minimal exposure to an unknown natural language. *Language Learning, 60,* 5-24. https://doi.org/10.1111/j.1467-9922.2010.00598.x

Hofweber, J., Aumônier, L., Janke, V., Gullberg, M., & Marshall, C.R. (2022). Breaking into language in a new modality: The role of input and individual differences in recognizing signs. *Frontiers in Psychology.* 13:895880. https://doi.org/10.3389/fpsyg.2022.895880

Hofweber, J., Aumônier, L., Janke, V., Gullberg, M., & Marshall, C.R. (2023). Which aspects of visual motivation aid the implicit learning of signs at first exposure? *Language Learning, 73,* Issue S1, 33-63. https://doi.org/10.1111/lang.12587

Hou, L., & Morford, J. (2020). Using signed language collocations to investigate acquisition: A commentary on Ambridge (2020). *First Language, 40,* 585-591. https://doi.org/10.1177/0142723720908075

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition, 27*(2), 129-140. https://doi.org/10.1017/S0272263105050084

Johnston, T., & Schembri, A. (2009). *Auslan: an introduction to sign language linguistics*. Cambridge University Press.

Klein, W. (1986). *Second language acquisition*. Cambridge University Press.

Lillo-Martin, D., & Hochgesang, J. (2022). Signed languages – ordinary and unique: A commentary on Kidd and Garcia (2022). *First Language, 42,* 789-803. https://doi.org/10.1177/01427237221098858

Mandel, M. (1981). *Phonotactics and morphophonology in American Sign Language*. PhD dissertation. University of California, Berkeley.

Meisel, J.M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, 3*(2), 104-135. https://doi.org/10.1017/S0272263100004137

Ortega, G., Schiefner, A., & Özyürek, A. (2019). Hearing non-signers use their gestures to predict iconic form-meaning mappings at first exposure to signs. *Cognition, 191,* 103996. https://doi.org/10.1016/j.cognition.2019.06.008

Perdue, C. (Ed.) (1984). *Second language acquisition by adult immigrants. A field manual.* Newbury House.

Rebuschat, P., & Williams, J.N. (2012). Implicit learning in second language acquisition. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. https://doi.org/10.1002/9781405198431.wbeal0529

Ristin-Kaufmann, N., & Gullberg, M. (2014). The effects of first exposure to an unknown language at different ages. *Bulletin Suisse de Linguistique Appliquée*, *99,* 17-29.

Sandler, W. (2012). The phonological organization of sign languages. *Language and Linguistics Compass, 6*(3), 162-182. https://doi.org/10.1002/lnc3.326.

Sandler, W., & Lillo-Martin, D. (2006). *Sign languages and linguistic universals*. Cambridge.

Schönström, K. (2021). Sign languages and second language acquisition research: An introduction. *Journal of the European Second Language Association,* 5(1), 30-43. https://doi.org/10.22599/jesla.73

Sockett, G. (2022). Learning beyond the classroom and autonomy. In H. Reinder, C. Lai, & P. Sundquist (Eds.), *The Routledge handbook of language learning and teaching beyond the classroom* (pp. 67-80). Routledge.

Wechsler, D. (1997). *WAIS-III Administration and scoring manual.* The Psychological Association.

Wechsler, D., Coalson, D.L., & Raiford, S.E. (2008). *WAIS-IV technical and interpretive manual.* Pearson.

Williams, J. (2009). Implicit learning in second language acquisition. In W. Ritchie & T. Bhatia (Eds.) *The new handbook of second language acquisition*. (pp. 319-353). Brill.

Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.

Data, Code and Materials Availability Statement

Data: available on the open science framework, https://osf.io/8hrp6/
Code: available on the open science framework, https://osf.io/8hrp6/, in the document 'R_scripts'
Materials: available on the open science framework, https://osf.io/8hrp6/, in the folder 'Experiment_regularities', in 'Weather forecast video' (accompanied by 'The weather forecast script in English'), and in the pdf document 'questionnaire'

The Editor granted exemptions to materials sharing for the following sets of materials, on the basis that they are proprietary and subject to copyright: (1) the English vocabulary subtest of the Wechsler Adult Intelligence Scale WAIS-IV (Wechsler et al., 2008), and (2) the matrices subtest of the Wechsler Adult Intelligence Scale WAIS-III (Wechsler, 1997).

Ethics Statement

This study was granted ethical approval by the Research Ethics Committee of the UCL Institute of Education, number REC 1156, on 18th January 2019. An amendment to allow for online data collection was approved by the same committee on 20th July 2020. All participants gave informed written consent before taking part in the study.

Authorship and Contributorship Statement

Julia Hofweber: conceptualization; data curation; formal analysis; methodology; investigation; visualization; writing – original draft preparation; writing – review and editing.
Lizzy Aumônier: conceptualization; formal analysis; investigation; methodology.
Vikki Janke: conceptualization; formal analysis; funding acquisition; methodology; project administration and supervision; writing – original draft preparation; writing – review and editing.
Marianne Gullberg: conceptualization; formal analysis; funding acquisition; methodology; project administration and supervision; resources; writing – original draft preparation; writing – review and editing.
Chloë Marshall: conceptualization; formal analysis; funding acquisition; methodology; investigation; project administration and supervision; resources; writing – original draft preparation; writing – review and editing.

Acknowledgements

This work was funded by an International Academic Fellowship (IAF2016023 awarded to Chloë Marshall) and a Research Grant (RPG-2018-333 awarded to Chloë Marshall, Marianne Gullberg, and Vikki Janke) from the Leverhulme Trust. We thank Linnéa Lindén, Peter Roslund, Henrik Garde and Josien Greidanus at Lund University Humanities Lab for their assistance with creating the experimental materials, and Vassilis Sideropoulos at the UCL Institute of Education for his statistical advice.