

Addressing Early Language Deprivation Through Interactive ASL Learning Technologies

Anonymous Author(s)
Institution Withheld for Review

Abstract

Language deprivation remains a critical challenge for Deaf and hard-of-hearing children when early exposure to sign language is limited. We present a web-based, interactive ASL learning environment designed to prioritize engagement through short, game-like lessons rather than static instruction. The system renders animated ASL using sign.mt and supports camera-based interaction via a custom computer-vision pipeline. To reinforce early exposure, we include a parallel caregiver interface displaying English alongside signs to support joint learning. We demonstrate the platform through interactive modules spanning math, stories, reading, and trivia highlighting its potential for scalable early ASL exposure.

CCS Concepts

• **Human-centered computing** → **Accessibility technologies**; *Collaborative interaction*; *Empirical studies in accessibility*; • **Applied computing** → **Education**; **Education of people with disabilities**.

Keywords

American Sign Language, language deprivation, Deaf and hard-of-hearing learners, disability equity, inclusive education, early childhood learning, accessibility technologies, camera-based interaction

ACM Reference Format:

Anonymous Author(s). 2026. Addressing Early Language Deprivation Through Interactive ASL Learning Technologies. In *Proceedings of the 2026 Conference for Research on Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT 2026)*, June 2026, Chicago, IL, USA. ACM, New York, NY, USA, 4 pages.

1 Introduction

Language deprivation occurs when a child lacks sufficient access to a fully accessible natural language during early development. Unlike delayed vocabulary growth, it reflects reduced exposure to structured linguistic input during the brain’s critical language-forming period [1,2], leading to measurable cognitive and neural differences.

In the United States, 2–3 per 1,000 children are born with detectable hearing loss [3], and over 90 percent are born to hearing parents [4]. Because most hearing parents do not already know

ASL, many children lack fully accessible linguistic input during infancy, increasing risk of early language deprivation.

Parents play a foundational role in early neurodevelopment. Everyday interactions support word–object mapping and symbolic learning [5]. During the first five years, the brain forms hundreds of new neural connections per second [6], and consistent accessible language stabilizes circuits supporting syntax and higher-order cognition; without it, these systems reorganize differently [7,2].

The brain does not prefer spoken over signed language. When ASL is acquired early, it recruits the same left-hemisphere language networks as speech [8]; accessibility and consistency—not modality—determine typical development.

However, while online ASL resources exist, most current solutions are static, text-heavy, or adult-oriented. Few platforms provide a seamless, interactive environment that simultaneously supports (1) parent learning, (2) child engagement, and (3) consistent visual–linguistic mapping in a developmentally appropriate format.

Animated signs are paired with English equivalents to reinforce sign–word correspondence and enable joint caregiver learning. The experience is designed to feel game-like while embedding structured linguistic reinforcement.

Given the long-term consequences of early language deprivation, structured ASL exposure in the home is a critical intervention.

2 Theoretical Framework

2.1 Experience-Dependent Development of the Language Network

Neuroimaging shows early language exposure shapes the structural and functional organization of the brain’s language network, particularly the arcuate fasciculus within the dorsal stream [9].

When early language exposure is delayed, measurable neural differences emerge. Deaf individuals who experience delayed access to a first language show reduced activation and altered connectivity within classical left-hemisphere language regions [7]. Diffusion tensor imaging reveals decreased white matter integrity in the arcuate fasciculus among individuals with language deprivation, with degree of alteration predicting language proficiency [10].

2.2 Sensitive Periods and Synaptic Plasticity

These differences emerge during a highly plastic developmental window, when neural circuits rapidly form and prune in response to environmental input, reaching up to 700–1,000 new synapses per second in early life [6].

Consistent linguistic mapping reinforces fronto-temporal pathways, whereas absent or inconsistent input alters canonical circuitry [7,10]. Early exposure is therefore architecturally formative.

This developmental principle is illustrated in Figure 2. From birth, infants possess highly active visual systems; the occipital cortex

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

RESPECT 2026, Chicago, IL, USA

© 2026 Copyright held by the owner/author(s).

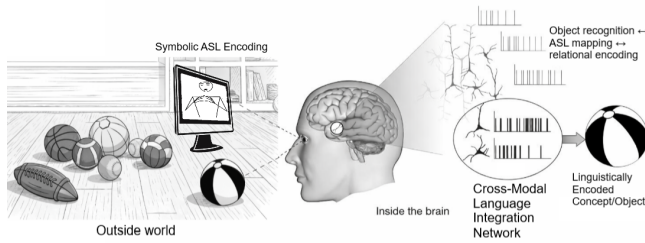


Figure 1: Early language exposure supports the formation of cross-modal integration networks within the brain. Consistent ASL or spoken language input strengthens fronto-temporal connectivity (e.g., arcuate fasciculus) within the dorsal language stream. When symbolic input is present, visual perception becomes integrated within a linguistically encoded conceptual framework.

rapidly organizes around processing objects, motion, contrast, and spatial relationships. What is not innate, however, is the stable linkage between a visual percept and a linguistic symbol. That association must be learned through repeated, structured exposure.

2.3 Symbolic Mapping: From Perception to Language

Without consistent linguistic pairing, object remains primarily perceptual, and coupling between perceptual and linguistic systems is weaker or reorganized [7].

In Figure 2, the distinction is not exposure to the object itself, but exposure to consistent, accessible linguistic labeling. Child A (hearing) repeatedly experiences contingent pairing between the visual object and a stable language label (“ball”), strengthening cross-regional connectivity between perceptual and language networks. Over time, the object becomes encoded not only as a visual form but as a symbolic category embedded within language.

In contrast, Child B (hard of hearing) perceives the same object but without reliable linguistic reinforcement. The perceptual representation may remain intact, yet the absence of structured mapping limits stable integration with canonical language circuitry. The difference illustrated is therefore not sensory perception, but the presence or absence of reinforced symbolic association during a period of heightened neural plasticity.

Developmental research describes this process as word-object mapping [13,5], whereby structured input shapes perceptual systems toward language-relevant categories. Early-acquired sign language recruits classical left-hemisphere language regions [8], underscoring that accessibility—not modality—drives typical development.

When early access is absent, measurable differences emerge in dorsal stream integrity [10], syntactic efficiency [9], and long-term functional organization [7]. Early bridging of perception and language is therefore foundational.

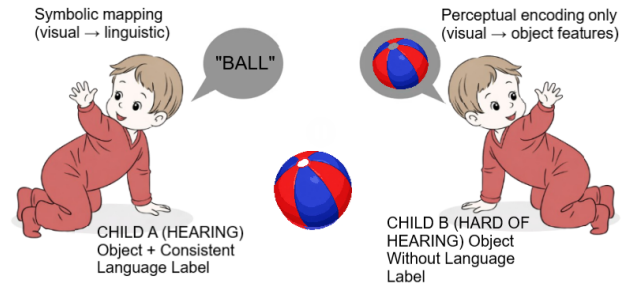


Figure 2: Symbolic mapping during early development. Left: Child A (hearing) receives consistent linguistic labeling paired with visual perception (e.g., “ball”), strengthening associative pathways between perceptual and language systems. Right: Child B (hard of hearing) perceives the same object without consistent linguistic reinforcement, resulting in perceptual encoding without stable symbolic integration.

2.4 Existing Work and Positioning

Existing ASL learning tools generally fall into two categories: (1) signed storybooks or avatar-rendered content that provide narrated ASL exposure paired with text or images, and (2) gamified applications that use camera-based recognition for isolated alphabet or vocabulary practice. The former are typically one-directional, while the latter emphasize discrete sign production rather than curriculum-structured or narrative lesson flow.

Research prototypes have explored narrative interaction combined with sign recognition, but these systems are often experimental and not publicly deployed for early-childhood structured learning.

To our knowledge, no publicly available camera-based ASL learning platform integrates ASL-first narrated instruction, child-oriented interaction loops with built-in performance pauses, and real-time computer-vision feedback within a unified system designed for joint child-caregiver learning.

3 System Design

3.1 System Overview

The system integrates animated ASL rendering with real-time camera-based interaction to support early language learning. Lesson scripts trigger sign.mt avatar playback, synchronized subtitles, and paired visuals for joint child-caregiver exposure. When interaction is required, a webcam-based verification loop confirms user responses and advances lesson progression.

Lesson scripts index structured content lines (LINES[idx]) that trigger sign.mt playback and synchronized UI updates. The architecture is intentionally modular: sign rendering, UI presentation, and interaction verification operate as distinct components connected through event-based state transitions.

3.2 Camera-Based Interaction

To enable interactive verification of user responses, the system incorporates a real-time hand tracking pipeline built on MediaPipe

Hands. As shown in Figure 3, a webcam stream is processed to extract 21 hand landmarks per frame. These landmarks are converted into a simple, interpretable finger-state estimate, including a thumb-distance proxy to differentiate ambiguous configurations.

A 500ms hold-to-confirm gate reduces transient misclassifications; only stable matches trigger feedback and progression. This lightweight geometric approach prioritizes interpretability and low-latency browser deployment over heavy classifier training.

3.3 Animated ASL Rendering via sign.mt

ASL content is rendered using sign.mt, which generates animated avatar-based sign language sequences. Each lesson line triggers a pre-rendered avatar video synchronized with subtitles and contextual visuals in the caregiver-facing UI. This dual-channel presentation reinforces sign-word correspondence and supports repeated symbolic mapping during early exposure.

The rendering pipeline is decoupled from the interaction loop, allowing the system to operate in two modes: passive exposure (continuous animated signing with UI reinforcement) and interactive verification (camera-based confirmation). This separation ensures consistent ASL exposure while preserving responsiveness during active tasks.

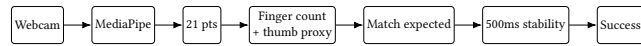


Figure 3: Condensed camera-based ASL interaction pipeline. A webcam stream is processed by MediaPipe Hands to extract 21 landmarks, converted into a simple finger-state estimate. A 500ms hold-to-confirm gate stabilizes predictions before emitting a success event.

4 Example Learning Modules

We present three module types for early language learners: (1) basic math, (2) spelling and vocabulary, and (3) short narrative stories.

Given that language deprivation occurs early, the modules emphasize simple, visually grounded curriculum reinforcing structured visual-linguistic pairing during peak neural plasticity (Sections 2.2–2.3).

4.1 Basic Math

The basic math module introduces small-number arithmetic (e.g., 1–5) using animated ASL number signs paired with visual object groupings. When a number is signed, the corresponding quantity is displayed visually, and users may be prompted to reproduce the sign via camera-based interaction.

This pairing reinforces sign-quantity-symbol mapping while strengthening ASL exposure.

4.2 Spelling and Vocabulary

The spelling and vocabulary module focuses on individual letters and simple, high-frequency words. Each letter is rendered via animated ASL fingerspelling and paired with a corresponding visual referent (e.g., an object image). Interactive prompts allow users to reproduce handshapes, reinforcing motor-symbol-concept associations.

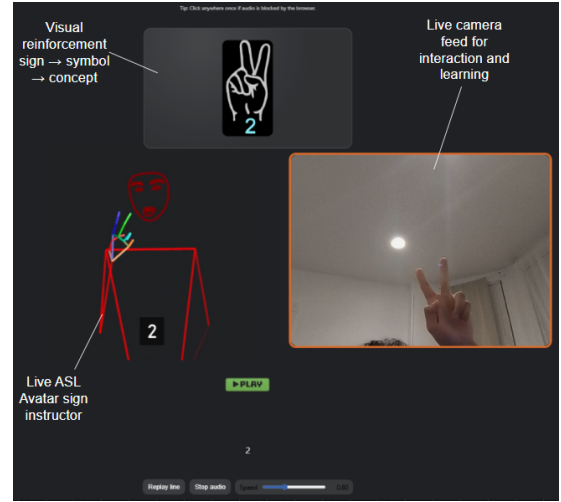


Figure 4: Basic math module. Animated ASL number signs are paired with visual quantity representations and optional camera-based interaction, reinforcing sign-symbol-concept mapping. Labels, in all modules, are illustrative overlays for the paper and are not part of the deployed user interface.

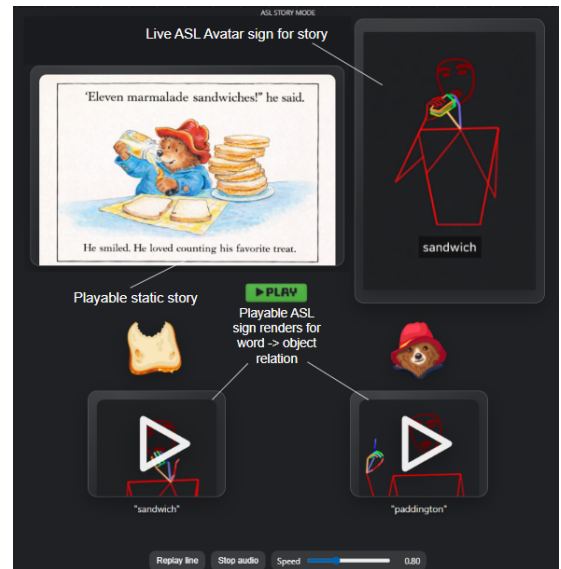


Figure 5: Story module. Narrative ASL rendering is synchronized with contextual visuals, supporting structured symbolic mapping across events and relational language.

This reinforces word-object mapping during sensitive developmental periods.

4.3 Reading Comprehension and Stories

Short, visually driven story segments provide continuous ASL exposure in narrative form. Animated signing is synchronized with

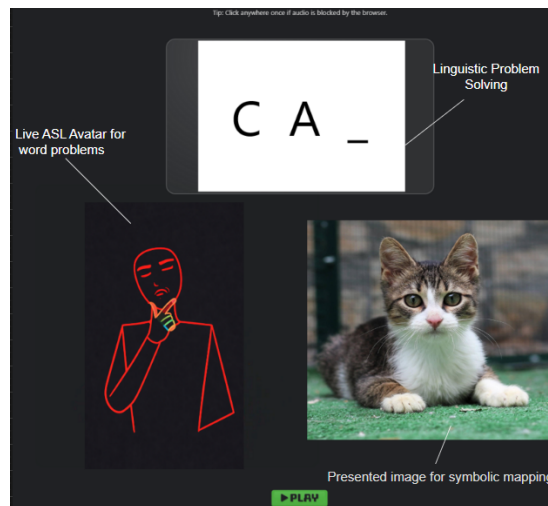


Figure 6: Spelling and vocabulary module. Animated ASL signs are paired with static images to strengthen word-object mapping and reinforce symbolic associations.

contextual imagery in the caregiver-facing UI, allowing both child and parent to associate signed language with unfolding events.

Narratives extend symbolic mapping beyond isolated labels to relational sequences, supporting early syntactic reasoning and sustained exposure within meaningful contexts.

Across modules, the emphasis remains structured pairing between perception and accessible language during peak neural plasticity.

5 Discussion and Limitations

The current prototype focuses on foundational early language elements, including numbers, basic vocabulary, and simple narratives. Camera-based interaction is limited to letters and numbers using lightweight geometric estimation rather than full sign recognition.

The avatar does not yet capture nuanced facial expressions and non-manual markers, which are linguistically meaningful in ASL. Additionally, regional variation and co-articulation are not fully modeled. These constraints limit phrase-level interaction and advanced grammatical fidelity.

The system is positioned as an early-intervention tool focused on structured symbolic mapping during sensitive developmental periods.

6 Conclusion and Future Work

Language deprivation during early childhood has measurable cognitive and neural consequences. By integrating animated ASL rendering with lightweight camera-based interaction, this work presents a scalable approach to reinforcing early symbolic mapping between perception and language.

Future work will expand interaction beyond letters and numbers to phrase-level recognition and richer avatar expressivity, particularly non-manual grammatical markers. Formal user studies with early learners and caregivers will evaluate developmental impact and usability.

This work aims to mitigate language deprivation at its developmental root through early, accessible ASL exposure.

References

- [1] Mayberry, R. I. Early language acquisition and adult language ability: What sign language reveals about the critical period for language. *Journal of Deaf Studies and Deaf Education*, 15(3):242–258, 2010. <https://doi.org/10.1093/deafed/enq005>
- [2] Hall, W. C., Levin, L. L., and Anderson, M. L. Language deprivation syndrome: A possible neurodevelopmental disorder with sociocultural origins. *Social Psychiatry and Psychiatric Epidemiology*, 52(6):761–776, 2017. <https://doi.org/10.1007/s00127-017-1351-7>
- [3] Centers for Disease Control and Prevention (CDC). Hearing loss in children, 2023. Available at: <https://www.cdc.gov/ncbddd/hearingloss/>
- [4] Mitchell, R. E. and Karchmer, M. A. Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, 4(2):138–163, 2004. <https://doi.org/10.1353/sls.2004.0005>
- [5] Kuhl, P. K. Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11):831–843, 2004. <https://doi.org/10.1038/nrn1533>
- [6] Center on the Developing Child, Harvard University. From best practices to breakthrough impacts: A science-based approach to building a more promising future for young children and families, 2016. Available at: <https://developingchild.harvard.edu/>
- [7] Mayberry, R. I., Chen, J.-K., Witcher, P., and Klein, D. Age of acquisition effects on the functional organization of language in the adult brain. *Brain and Language*, 119(1):16–29, 2011. <https://doi.org/10.1016/j.bandl.2011.05.007>
- [8] Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D., and Evans, A. C. Speech-like cerebral activity in profoundly deaf people processing signed languages: Implications for the neural basis of human language. *Proceedings of the National Academy of Sciences*, 97(25):13961–13966, 2000. <https://doi.org/10.1073/pnas.97.25.13961>
- [9] Skeide, M. A., Brauer, J., and Friederici, A. D. Brain functional and structural predictors of language performance. *Cerebral Cortex*, 26(5):2127–2139, 2016. <https://doi.org/10.1093/cercor/bhv042>
- [10] Cheng, Q., Roth, A., Halgren, E., and Mayberry, R. I. White matter alterations in children with language deprivation. *NeuroImage: Clinical*, 24:102085, 2019. <https://doi.org/10.1016/j.nicl.2019.102085>
- [11] Bavelier, D., Tomann, A., Hutton, C., Mitchell, T., Liu, G., Corina, D., and Neville, H. Visual attention to the periphery is enhanced in congenitally deaf individuals. *Journal of Neuroscience*, 21(17):RC93, 2001. <https://doi.org/10.1523/JNEUROSCI.21-17-j0001.2001>
- [12] Fine, I., Finney, E. M., Boynton, G. M., and Dobkins, K. R. Comparing the effects of auditory deprivation and sign language within the auditory and visual cortex. *Journal of Cognitive Neuroscience*, 17(10):1621–1637, 2005. <https://doi.org/10.1162/089892905774597173>
- [13] Werker, J. F. and Tees, R. C. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1):49–63, 1984. [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3)
- [14] Bond, M. *A Bear Called Paddington*. Illustrated by Peggy Fortnum. William Collins, London, 1958.