

Executive Summary

Lingo Lens is an augmented reality language learning application that transforms everyday surroundings into an interactive vocabulary builder. By leveraging Apple's Vision framework and translation services, the app allows users to point their camera at objects, identify them through object detection, and anchor translation labels in 3D space. Users can tap these labels to view translations and hear correct pronunciations in their target language, with the ability to save words for future reference in a personalized collection. The app addresses the challenge of contextual vocabulary acquisition by connecting words directly to physical objects, making language learning more intuitive and memorable than traditional flashcard methods. Technically, Lingo Lens combines ARKit for spatial tracking, Vision for object recognition, CoreML for inference, Apple's Translation framework for language services, AVFoundation for speech synthesis, and CoreData for persistent storage—all while functioning offline after initial language downloads, ensuring privacy and accessibility in various learning environments.