Do Ambient AI Scribes Propagate Interpreter Errors? A Pilot Study with Simulated Encounters

Alexandra Rabotin, MD, MPH[1], Efren Aguilar, BA[2], Saitiel Sandoval Gonzalez BS, MPH[2], Alonso Iniguez BS, MS[2], Douglas Bell MD, PhD[2], Christina Jung, MD, MPH[3-5], Eric Lee, MD[3], Jeffrey Arroyo, MD[3]

[1]Mission Community Hospital Family Medicine Residency Program, Panorama City, CA; [2]University of California Los Angeles Health Sciences, Los Angeles, CA; [3]AltaMed Health Center, Los Angeles, CA; [4]Children's Hospital Los Angeles, Los Angeles, CA; [5]Keck School of Medicine University of Southern California, Los Angeles, CA

Introduction

- Interpreter errors are common and can compromise care.[1-3]
- Ambient AI scribes are increasingly used, yet their performance in multilingual encounters is underexplored. It is unknown if they **propagate or mitigate** interpreter errors.
- This pilot study evaluated documentation of interpreter errors by AI scribes.

Methods

- Five simulated Spanish-English clinical scenarios with 20 deliberate interpreter errors (Table 1) were analyzed.
- Progress notes were generated by two commercially available AI scribes with multilingual comprehension capabilities and independently reviewed by two researchers.
- Each error was classified by type (omission, substitution, or addition) and by context (interpreting patient vs provider speech) (Figure 1a).
- Propagation was defined as the presence of the interpreter's error within the AI-generated documentation.

Table 1. *Examples of interpreter errors by type.*

Error type	Interpreter error
Omission	Patient: "He coughs at night and has wheezing" Interpreter: "He coughs at night" (wheezing omitted)
Substitution	Provider: "Take 2 tablets every 6 hours" Interpreter: "Take 2 tablets every 4 hours"
Addition	Provider: "Just humidifier and honey" Interpreter: "Just humidifier, honey, and a little lemon"

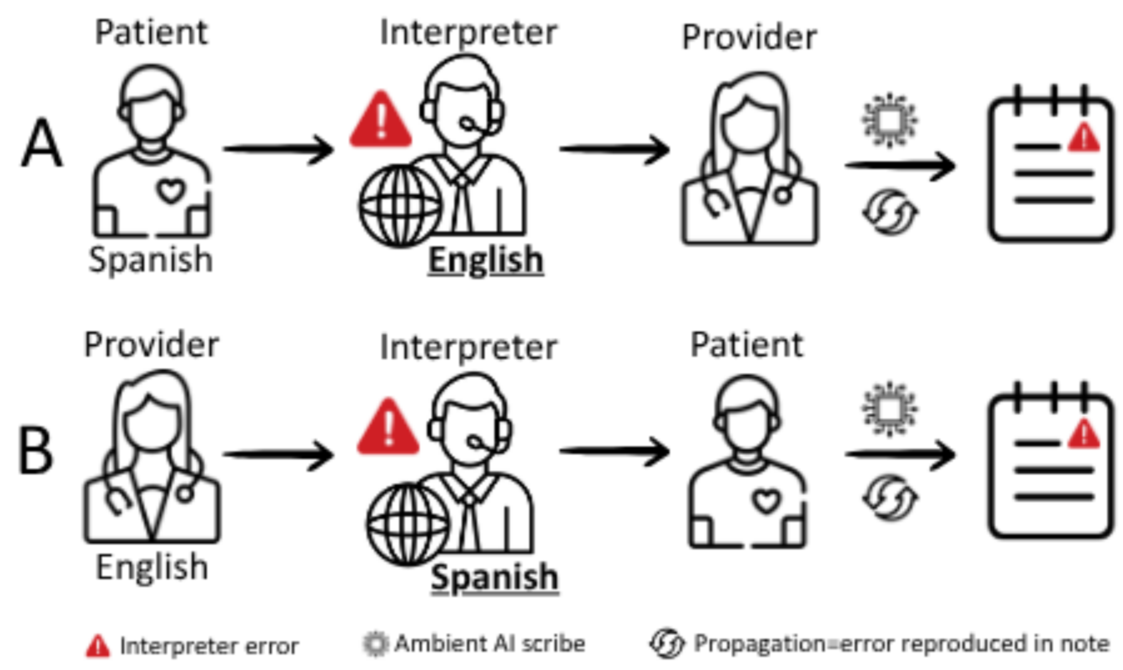

Figure 1a. *Context of interpreter error and propagation.* Errors from patient (A) or provider (B) speech may be reproduced in the clinical note.

Results

- Overall, 59% of interpreter error were propagated (range 25-89% depending on type) (Figure 1b).
- When interpreting patient speech, errors were incorporated in the majority of cases for both vendors. In contrast, when interpreting provider speech, propagation was far lower.

Discussion

- AI scribes frequently propagated interpreter errors, particularly when transcribing patient speech.
- This pattern likely reflects training bias toward English-dominant data.
- Implementing safeguards, such as discrepancy alerts or multilingual validation, may improve documentation safety in multilingual care.

References

1. Jackson JC, Nguyen D, Hu N, Harris R, Terasaki GS. Alterations in medical interpretation during routine primary care. J Gen Intern Med. 2011;26(3):259-264.
2. Pham K, Thornton JD, Engelberg RA, Jackson JC, Curtis JR. Alterations during medical interpretation of ICU family conferences that interfere with or enhance communication. Chest. 2008;134(1):109-116.
3. Flores G, Laws MB, Mayo SJ, et al. Errors in medical interpretation and their potential clinical consequences in pediatric encounters. Pediatrics. 2003;111(1):6-14.

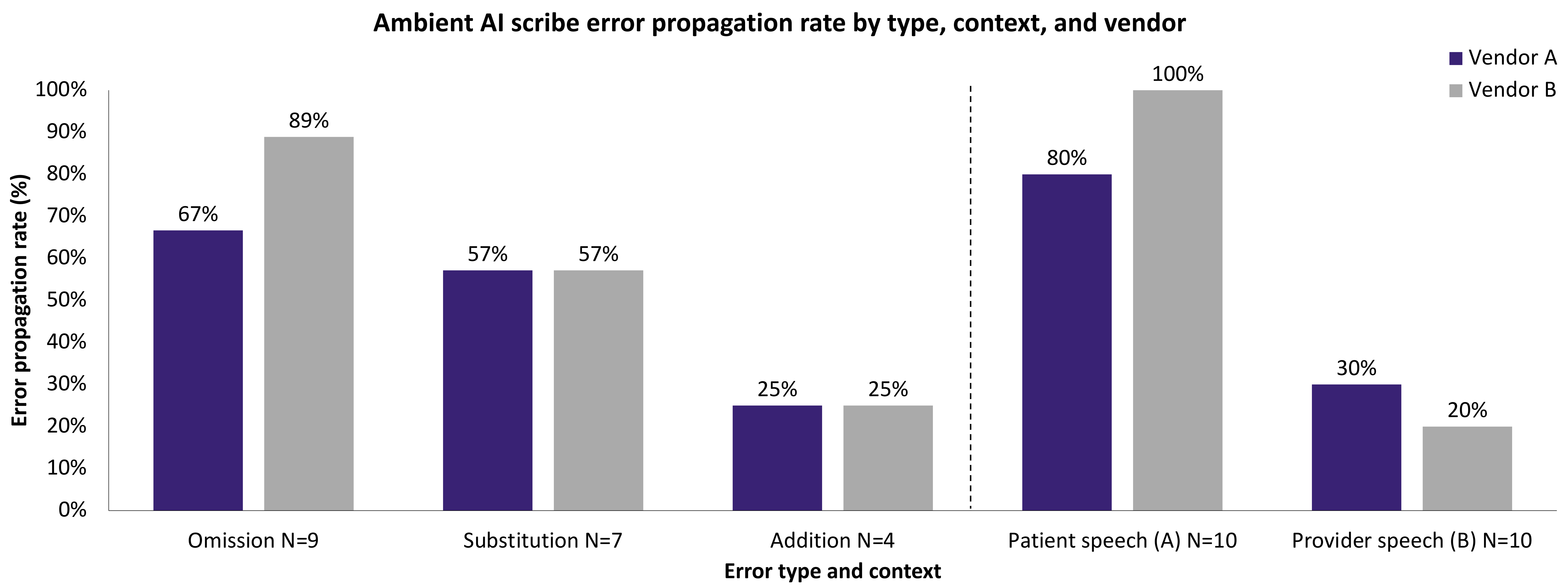

Figure 1b. *Ambient AI scribe error propagation rate by type, context, and vendor.* Error context: patient (A) vs provider (B) speech (see Figure 1a).

Ambient AI Scribes Can Propagate Interpreter Errors in Multilingual Care

Mission Community Hospital Compassionate Healthcare. Quality Healthcare. | AltaMed | UCLA