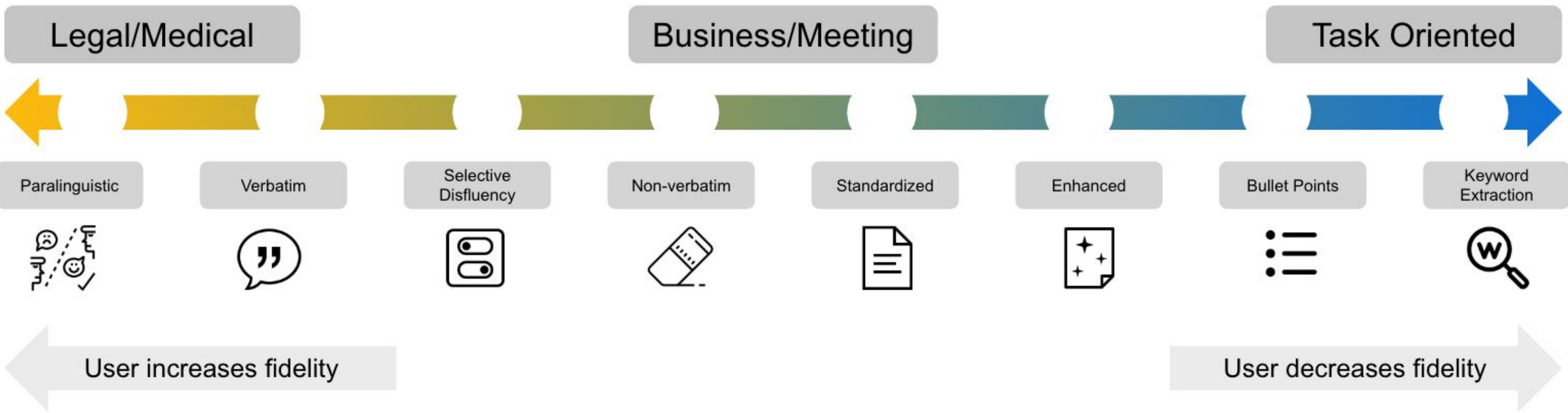
SpeechSpectrum

Anna Seo Gyeong Choi*, Maria Teleki*, Miguel del Rio, James Caverlee, Corey Miller, Allison Koenecke
Cornell University, Texas A&M University, Rev

Theoretical Contribution



The Linguistic Fidelity Spectrum, a spectrum of valid representations for the same speech.

Empirical Study

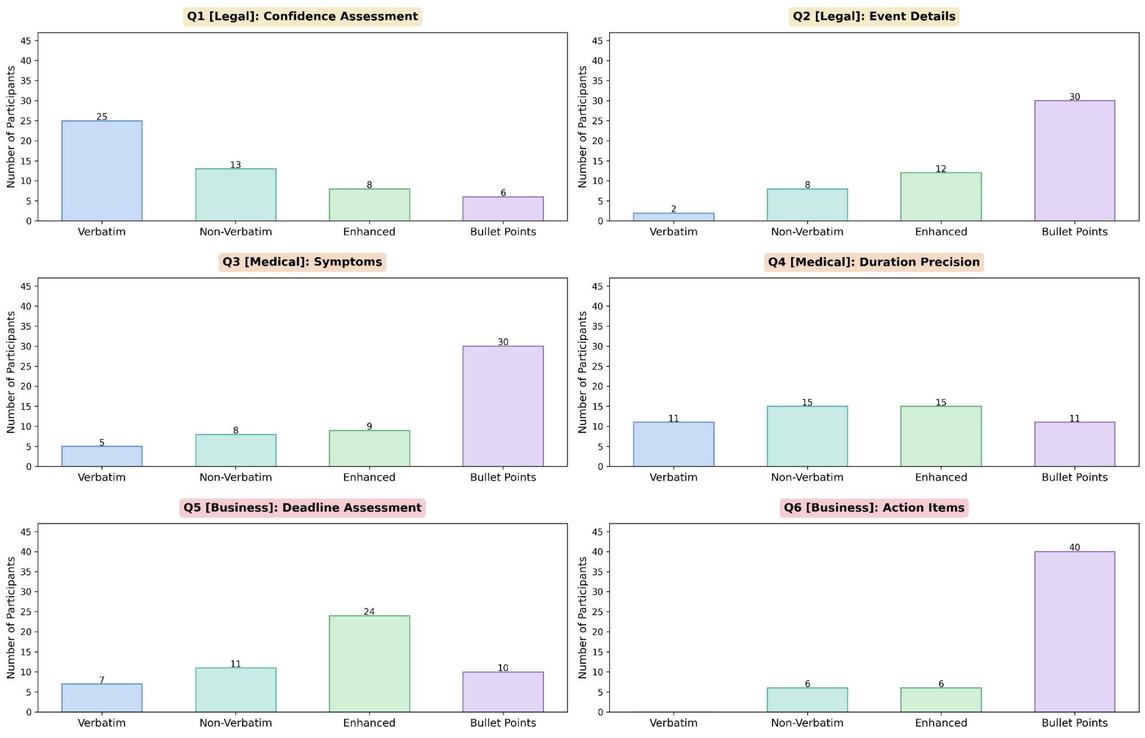
Legal Domain Imagine you are a case judge reading through a deposition transcript
Q1: Did the defendant seem confident about the details of the crash?
Q2: What were the events leading up to the crash?
Medical Domain Imagine you are a doctor looking over a triage dictation provided by a nurse
Q3: What are the main symptoms the patient is exhibiting?
Q4: Has the chest pain been going on for exactly three days, or could it have been longer/shorter?
Business Domain Imagine you are a team leader reading a meeting transcript
Q5: Does the team seem like they will meet the December deadline?
Q6: What are the action items from the meeting?

User Study (N=52). We ask questions spanning three domains, to determine if different domains beget different parts of the linguistic fidelity spectrum. For example, emotion assessment necessitates verbatim transcripts.

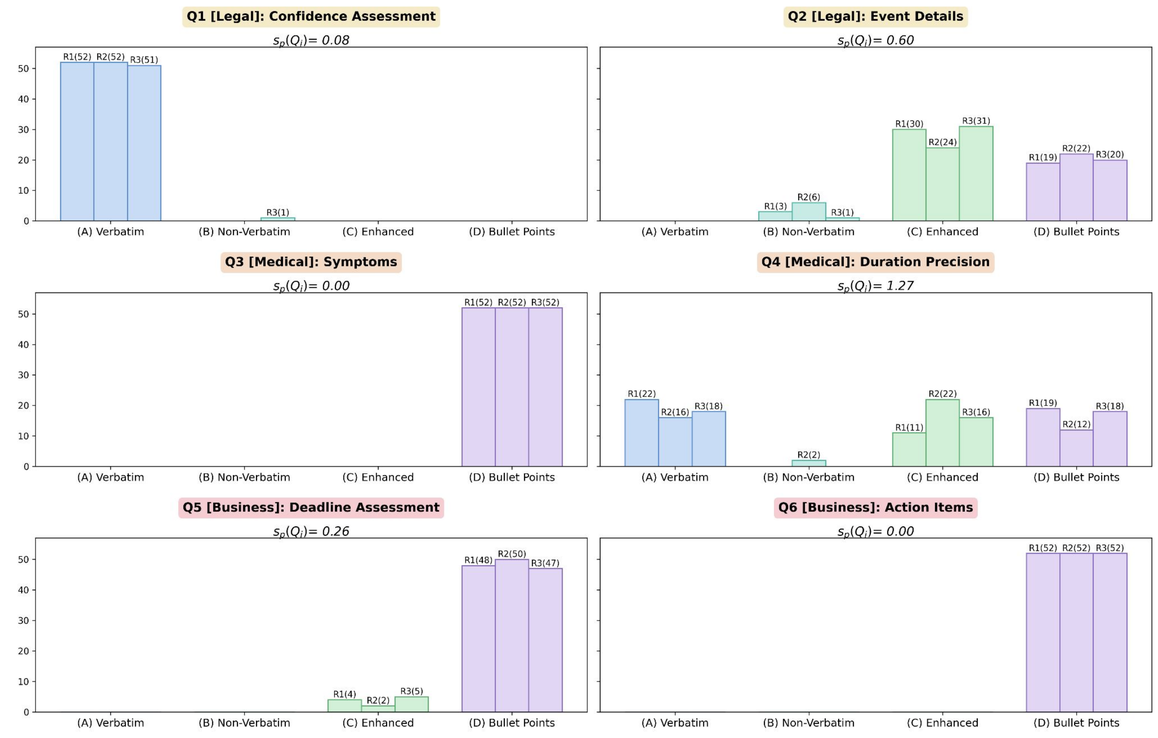
Respond as a person who **[(P1) does/does not]** work in automatic speech recognition technology, **[(P2) does/does not]** work in STEM (science, technology engineering, mathematics), and **[(P3) has legal expertise/has medical expertise/does not have legal or medical expertise]**. Respond only with the letter for the answer choice.

LLM Study (N=52 personas, 3 rounds). We prompt gpt-5.1 to act as the personas represented in our user study. We run each persona 3 times, for 3 temperatures.

Results



Human Preference Distribution. Some tasks show diffuse distributions, while others are sharper.



LLM Preference Distribution. LLMs rarely output diffuse distributions, regardless of the task.

Design Recommendations

- ▷ **R1: Support multi-fidelity interaction.**
- ▷ **R2: Incorporate task-aware defaults.**
- ▷ **R3: Prioritize task defaults over domain defaults.**
- ▷ **R4: Provide educational scaffolding.**