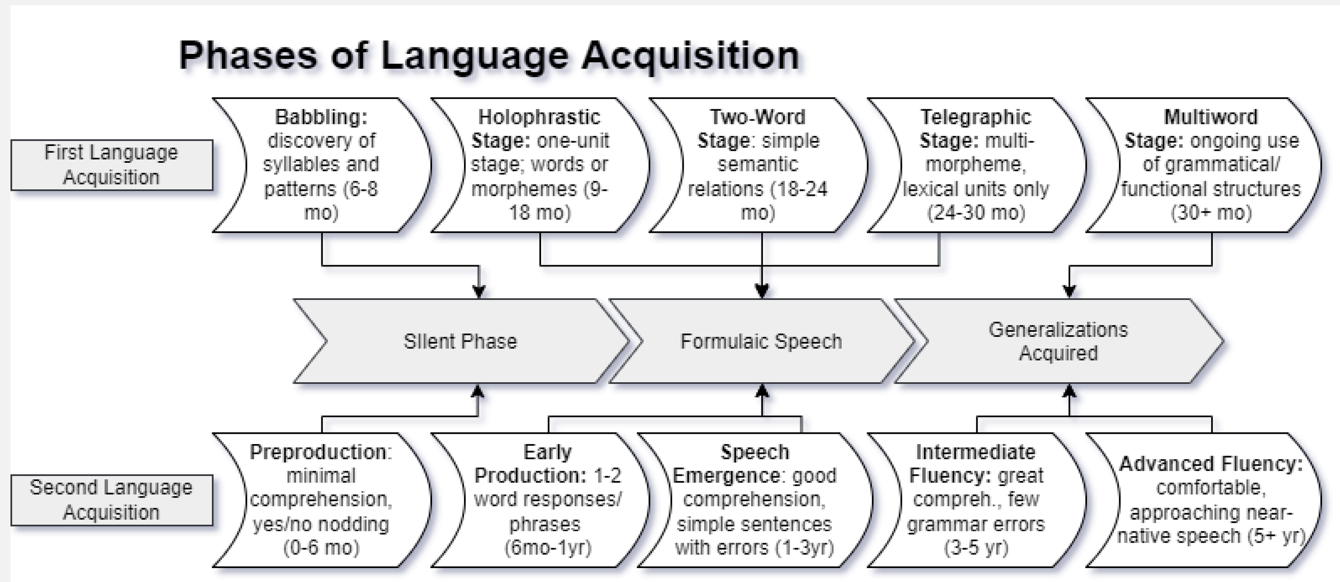
Creating Naturalistic Interactions in Language Learning Technologies

Brody Silva, Elizabeth Im, Demi Liu, Carol Lu, Anna Shinn, Bea Tran, Zihan Wang
Department of Cognitive Science, Johns Hopkins University
Research Mentors: Monica Lopez-Gonzalez

Background

Language pedagogy research tends to ignore many features identified in SLA as factors that determine success, such as stress/fear of failure, *i + 1* input, and communicatively embedded output. If the goal of language is successful communication, we must trigger learners to "notice the gap" in their production during real-time conversations, either in morphosyntactic (form) or semantic/pragmatic (meaning) errors.



Current Issues

Pedagogical Limitations

- The language taught through LLTs can be rigid, awkward, unsuitable for conversational use, and promotes the use of abbreviated forms.

Social Limitations

- LLTs lack an engaging factor, socialization, opportunities to practice in real-life contexts, and a necessary variety of supporting resources (both learning and technology related).
- The feedback provided by LLTs is generic.

While there are currently some examples of LLTs using VR (e.g. Mondly), these programs replicate the negative feedback learning of text-based websites. AR programs use place-based or marker-based methods to teach with real objects but rely on a limited set of known terms to function. Such a rigid structure prevents the user from receiving benefits of immersion. Instead, an immersive LLT must be (1) adaptive, (2) robust, and (3) output-driven.

	Babbel	Rosetta Stone	Duolingo
Real-life conversations*	✓	✓	✓
Vocabulary	✓	✓	✓
Grammar & syntax	✓	✓	✓
Reading	✓	✓	✓
Writing	✓	✓	✓
Listening (native speakers)	✓	✓	✓
Speaking (speech recognition)	✓	✓	
Repetition & Patterns	✓	✓	✓
Instant feedback	✓	✓	✓
Varied exposure to language	✓	✓	✓
Tips & explanations	✓		✓
Personalization of learning difficulty			✓

Example

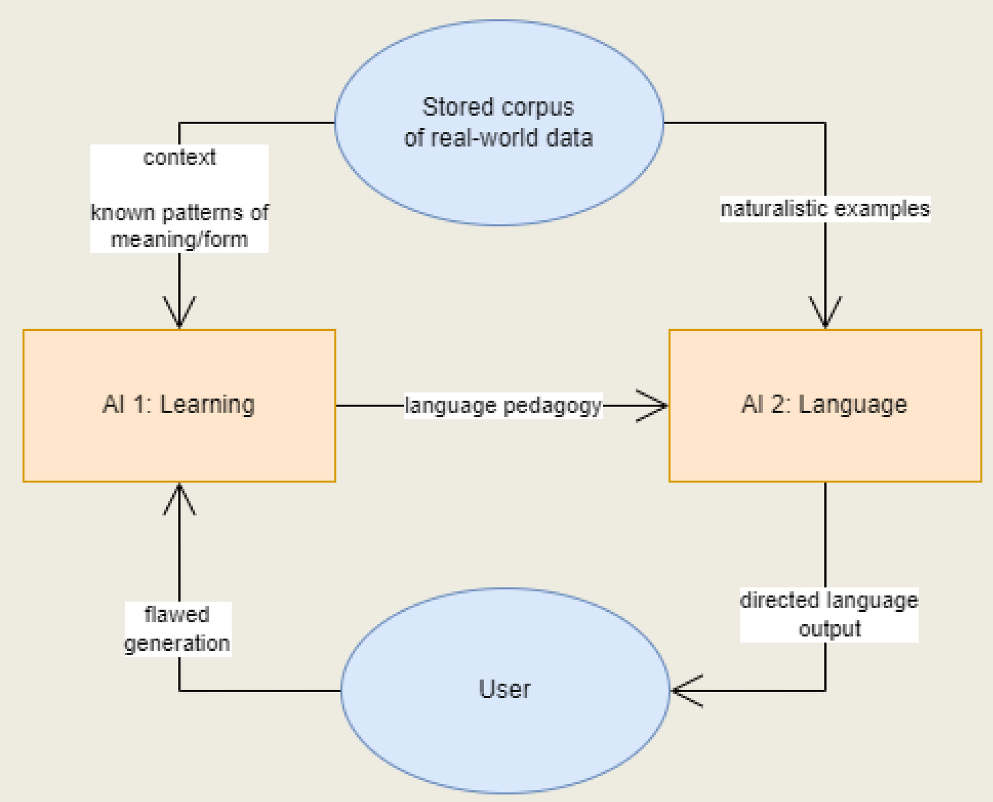
PREFUNCTION: Situation 1.a (Coffee shop. Ordering a drink)
(Italian): "***Vorresti qualcosa da bere***?" ("Do you want something to drink?")

Error of Meaning

- AI 1:
  - Input: "*Vorrei un cane.*" ("*I want dog.*")
    - [<1st.sing> Request <animal>]
  - Process: Identify error of meaning
    - ["*cane*" = <animal> when expecting <drink>; SEMANTIC]
  - Output: GOAL to use type expected.
    - [Request <drink> SEMANTIC]
- AI 2:
  - Input: GOAL = SEMANTIC <drink>
  - Process: Identify probable alternative [<drink> -> *caffè*]
  - Output: "*Vorresti un caffè o un cane*?" (*"Do you want à coffee or à dog?"*)

Error of Form

- AI 1:
  - Input: "*Vorresti un caffè."* ("You want coffee.")
    - [<2nd.sing> Request <drink>]
  - Process: Identify error of form
    - ["vorresti" = <2nd.sing> when expecting <1st.sing>; FORM]
  - Output: GOAL to use form expected.
    - [Request <1st.sing> FORM]
- AI 2:
  - Input: GOAL = FORM <1st.sing>
  - Process: Identify probable structure [<1st.sing> -> *vorrei*]
  - Output: "*Vuoi dire <<vorrei un cane>>?*" (*"Did you mean <<I want coffee>>?"*)



Proposal

To better simulate natural interactions, we propose a system of AI-generated language interactions coupled with VR/AR. By creating a goal-directed situation and a visible interlocutor, we artificially activate the necessary precursors to allow for language acquisition under Social Interactionist Theory (Gass, Mackey, & Pica 1998): goal-oriented communication, situations lacking in stress or anxiety, and dynamic conversation.

- We rely on **real-world data** of catalogued interactions sorted by **contexts** (e.g. coffee shop, library, transportation) and **conversations** (e.g. ordering coffee, asking for the restroom, mistaken order).
- We then use statistical models to predict expected responses in two ways: **forms** identified via n-gram models matched to predictable patterns of syntax/morphology; **meaning** semantic categories and approximate alternatives.

AI 1 identifies any errors in speech and determines a pedagogical goal which AI 2 realizes.

AI 1: Learning
Input: User speech (non-prosodic ASR)
Process: intelligibility rating; expected meaning/form (statistical corpus data)
Output: Goal (user to generate)

AI 2: Language
Input: Goal (error type + expectation
Process: Probable correction, naturalistic responses (statistical corpus data)
Output: Directed language

This process cycles throughout the conversation until the contextual goal is achieved.

Discussion

- The context-centric design allows for inductive learning through a combination of story-telling and direct engagement.
- This framework is theoretical; implementation requires a great deal more research on computational plausibility.
- Many LLTs show bias towards non-Romance languages with structures, pronunciations, and cultural properties (e.g. honorifics or registers) being difficult to encode. The proposal must be designed with the language's demands in mind to predict forms/meaning that are useful and important.
- The model assumes a simple process for speech-to-text recognition for linguistic processing, but this issue is confounded by dialect differences of input training and the non-native productions of users.

References

AppMagic. (November 15, 2021). Leading language learning apps worldwide in October 2021, by downloads (in 1,000s) [Graph]. In Statista. Retrieved December 12, 2021, from https://www.statista.com/statistics/1239522/top-language-learning-apps-downloads/

Babbel. (19 November, 2018). *An Expert Explains: The Babbel Approach*. [Video]. YouTube. https://www.youtube.com/watch?v=7sr12tqDZNA

Babbel. (n.d.). *The babbel method*. Babbel. https://about.babbel.com/en/.

Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, *23*(4), 459-484.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623).

Brown, H. D. (2007). Principle of learning and teaching. New York, NY: Pearson Education.

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. in J. R. Hayes (ed.), cognition and the development of language (pp. 11-53). New York Wiley.

Corder, S. P. (1967). The significance of learner's errors.

Dulay, H. C., & Burt, M. K. (1974). Natural sequences in child second language acquisition 1. *Language learning*, *24*(1), 37-53.

Duolingo. (n.d.) *About Us.* Duolingo research. https://research.duolingo.com/.

Duolingo. (n.d.) Our teaching approach. Duolingo. https://www.duolingo.com/approach.

Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. *The handbook of second language acquisition*, *14*, 63.

Ellis, R. (1984). *Classroom second language development: A study of classroom interaction and language acquisition*. Pergamon.

Ellis, R. (1997). SLA and language pedagogy: An educational perspective. *Studies in Second language acquisition*, *19*(1), 69-92.

Gass, S. M., Mackey, A., & Pica, T. (1998). The role of input and interaction in second language acquisition: Introduction to the special issue. Modern Language Journal, 299-307.

Khajavy, G. H., MacIntyre, P. D., & Barabadi, E. (2018). Role of the emotions and classroom environment in willingness to communicate: Applying doubly latent multilevel analysis in second language acquisition research. *Studies in Second Language Acquisition*, *40*(3), 605-624.

Krashen, S. (1982). Principles and practice in second language acquisition.

Loewen, S., Crowther, D., Isbell, D., Kim, K., Maloney, J., Miller, Z., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL*, *31*(3), 293-311. https://doi.org/10.1017/S0958344019000065

Loewen, S., Isbell, D. R., & Sporn, Z. (2020). "The effectiveness of app-based language instruction for developing receptive linguistic knowledge and oral communicative ability. *Foreign Language Annals*, *53*, 209-233. https://doi.org/10.1111/flan.12454

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language teaching*, *51*(3), 285-329.

Long, M. H. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Sciences*, *379,* 259–278. https://doi.org/10.1111/j.1749-6632.1981.tb42014.x

Lord, G. (2015). "I don't know how to use words in Spanish": "Rosetta Stone" and Learner Proficiency Outcomes. *The Modern Language Journal*, *99*(2), 401–405. http://www.jstor.org/stable/43650036

Lord, G. (2016). "Rosetta Stone for Language Learning: An Exploratory Study". *IALLT Journal of Language Learning Technologies*, *46*(1), 1-35, https://doi.org/10.17161/iallt.v46i1.8552.

Ma, Q. (2017). A multi-case study of university students' language-learning experience mediated by mobile technologies: A socio-cultural perspective. *Computer Assisted Language Learning*, *30*(3-4), 183–203. https://doi.org/10.1080/09588221.2017.1301957

Acknowledgements

Thank you to the Johns Hopkins Federal Credit Union and Alumni Association for sponsoring this poster; to Julia Yarmolinskaya for providing background and education about language acquisition; and to Monica Lopez-Gonzalez for advising and critiquing this work as it continues.