

# Target Audience: You

*In linguistics, the study of idiolects explores the individual's dialect at one time. The definition of the word idiolect inherently indicates that everyone uses language differently and that each person's use of a language constantly changes. This project explores an individual's relationship with their idiolect and how to optimize their use. We use the word optimize as an umbrella term for feeling empowered by one's idiolect, confident in it, and responsible in using it. We put its components through Natural Language Processing's back translation-inspired approach to learn about idiolects from diverse lenses. The process maps the elements of an idiolect into the dictionary of a non-linguistic field (e.g., computer science, visual arts), studies it using the diction of said "foreign" field, and finally translates it back to linguistic-friendly diction. NLP uses this approach to augment data, but this project hopes to use back translation to uncover hidden/new perspectives.*

*Moreover, the choice of contrasting fields like computer science and the arts hopes to humanize an otherwise non-human exploration, given the fellow's AI Language Model-heavy view of language. We focus on the idiolect's prominence in an individual's self-dialogue (within the mind) and with others. We also scale our focus to the English language idiolect of an international ESL (English as a second language) student. Although the exploration is ongoing beyond the temporal scope of the fellowship, we find the identity and self as prominent grounding variables of the idiolect.*

## The Beginning

Despite my past dreams of being a writer, storyteller, and productive politician, all of which I believe requires a level of expertise and control over one's language, 3.5 years of purely Math and Computer Science courses resulted in the breaking of my sense of control and expertise over what comes out of my mouth. I had the materials needed to write an essay criticizing the way I speak, and that's not really health is it? My desire to outgrow these (self-imposed?) chains motivated this exploration of idiolects: the language peculiar to a particular person. Idiolects change over time, so would it be possible to make use of my 3.5 years of loss (in the context of my language) that is CS and Math (and with that, Artificial Intelligence) to learn about and change my own idiolect so that I can tell anyone anything in any amount of time?

## Chatbot

The beginning of my exploration was a desire to build some system to model and experiment with idiolects without much human interference. This became an attempt of making chatbots with its own idiolect that would facilitate experiments where a bot would be tasked to explain an idea to a human in a finite amount of time. Then, the produced conversation would be studied at two levels. We'd first ask whether the human understood the idea the bot was trying to explain and questions around how they felt during and after the conversation. Then, we'd show the bot's idiolect to the human, and ask if their opinions change (ie. if you knew that the entity you're talking to has a habit of asking pushing questions, how would your overall outlook of a conversation with it change?). The bot would run multiple experiments where the idea to be explained is fixed, and the amount of time given varies, in order to answer whether or not it is possible to tell anyone anything in any (sensible) amount of time.

To enable the dissecting of the bot's idiolect, one fundamental idea was reinforcement learning (RL), a subfield of AI that relies on optimizing a numerical value measured using a "reward function". I thought these functions could comprise of the different aspects of an idiolect like percentage of positive words used, word length, number of questions asked, difficulty of words used, etc. However, I soon found that the merging of RL and Large Language Models is complex and time-consuming for my level of undergraduate AI knowledge... The implementation simply did not work, and no research papers thus far was able to help me find a solution feasible within 5 weeks. For now, this chatbot remains on pause.

## **Mirrors, LLM Training**

Although the chatbot dream lacked fruition, the AI-lens gave me a perspective of where the idiolect makes individuals stand in a larger group. In *The Mirror of the Mind*, Chomsky talks of language as a mirror reflecting complex systems like society, culture, and creativity of the mind. If we narrow down "language" to "individual language", and therefore idiolect, we then have a reflection of one person's experience of the world up to a point in time. This sounds a lot like the pre-training process of a language model (LM) in NLP, a stage where a mathematical model would be fed large amounts of text data and learn probabilities. This is done such that by the end of training, feeding the LM the phrase "I eat" would result in the LM predicting higher probabilities for the word "apples", because it is edible, rather than "sand".

Through this lens, the text data used in training is equivalent to the accumulation of life one person's lived at a time, while the prediction of the most probable next word becomes a watered down version how we actually produce every day language. There are many other fun mappings between LM components and our use of idiolects, but none of them as interesting to me, as the significance of LM "fine-tuning". See, while the pre-training of an LM allows it to get a general probability understanding of a language, the feeding of domain-specific data is often needed for the LM to be used in real life. For example, an LM used for a healthcare chatbot would need to be fine-tuned on healthcare data in order to increase the probability of it producing language in the appropriate lingo. As humans, we do the same automatically when talk differently at work compared to at home. But more importantly, the same way LMs need pre-training and fine-tuning, our idiolects play a balancing game between averaging and maintaining uniqueness. Our idiolects need to be able to produce an average language that maps ideas correctly to other people's idiolects, but as mirrors of our identities, the idiolects need to be different enough so that a close friend can still say: "yeah, that's something only [your name] would say".

The previous realization brought a halt to my CS/AI oriented view of things. This is because from the get go of this realization, I noticed that much of the balancing between averaging one's idiolect into the world's (a matter of being understood, and transferring information) and preserving uniqueness (keeping this mirror an authentic reflection of one's identity and life experiences) seem to be automatic and natural rather than formulated. I spent the latter third of my project exploring this idea, and I'll spend the rest of this write up posing guiding questions that helped me figure out where I stand in between the two ends, and I hope it helps you.



## What does the voice in your head mean to you?

*(Documenting Uniqueness)*

**What to do:** Talk to yourself silently and once you're immersed, grab a pen and write out the voice in your head, or record yourself out loud. Make a note of how you feel about what you've produced, right after finishing. This is a personal record meant to be kept for yourself, so remember that there are no expectations, no evaluation.

**Use it to:** Keep a record of how your idiolect changes over time, and remember peculiarities about your idiolect that you like. In this case, it would be good to review the product shortly after, and identify peculiar parts of your language that you know the reasoning behind.



## Would you share the voice in your head?

*(Identifying how you average)*

**What to do:** After writing/recording enough of these entries, find one that you like the most. Would you share this with someone?

**Use it to:** For those who are not confident in their idiolect, to narrow down what part of their interaction with their idiolect they can spend time on. Look at the note on how you felt after making the product. The dynamic between the feeling note, and whether or not you'd share your product may inform how you feel about your idiolect. If you could edit your product, would you share with different people? What kinds of filters did you use in these edits, and are they motivated by your relationship with the people you would now share it with?



## Who was coming after you, and who are they today?

*(Guiding your idiolect)*

**What to do:** A reflection of times you regretted saying something, remembering the parties of that conversation, and what worries you had that made you regret.

**Use it to:** Direct your idiolect. Were you able to remember any of these moments, and how much value do you put into that worry today? Who were the people in these memories of regret, and do they still matter to you today? Was the regret a result of external reaction, or sourced internally? For the last question, asking yourself whether you'd say the exact same thing to a different person might help. Based on your answers, reflect upon the components of your idiolect that caused you worries (if any), and ideate goals on change. You can flip the negative sentiment words in this section for positive ones, for reflection upon what components of your idiolect you'd like to preserve.

## **Some Notes:**

I spent most of my life around people that value anything tech, business, engineering, and innovation, all of which made me an incredibly product-oriented person. At some point, I felt that I needed to make a chatbot, although I knew for a fact that it was no longer as valuable to me as learning more about my own relationship with my idiolect, and stumbling upon with questions surrounding it that would really keep me up at night. While I felt engaged thinking about other people's questions, the excitement left me when I returned to my question of "how do I make this chatbot work?". It took me weeks, and a scary declaration of no longer wanting to make a product, until I was finally able to retrace my true personal motivation behind the project (ie. I don't like the way I speak) and find questions around idiolect that I genuinely felt excited to answer (these are predecessor questions to the ones I posed above). I don't think this "insight" is for everyone, but for those who relate to me, and feel guilt whenever they're not working on "something that would impact communities", I think you can always consider tracing back to the start, and using your reactive excitement to a question as a gauge of whether or not to pursue it.



Evaluasi (Reprise)  
Hindia



\*the song that ended my years of worrying over my own idiolect.



UP NEXT

LYRICS

RELATED

Yang...  
Tak bisa terobati  
Biarlah...  
Mengering sendiri, Menghias tubuh dan

Yang...  
Mengevaluasi ragamu  
Hanya kau sendiri

Mereka tak mampu melewati  
Yang telah kau lewat  
Tiap berganti hari  
Rintangannya yang kau hadapi

Masalah yang mengeruh (uu-uu)  
Oh perasaan yang rapuh  
Ini belum separuhnya  
Biasa saja, kamu tak apa  
Yang...

Selalu ingin ambil peran  
Hanya berlomba  
Menjadi lebih sedih dari dirimu

Muak dikesampingkan, disamakan  
Hatimu terluka, sempurna

Masalah yang mengeruh  
Oh, perasaan yang rapuh  
Ini belum separuhnya <sup>A</sup>  
Biasa saja, kamu tak apa

Perjalanan yang jauh  
Kau bangun untuk bertaruh  
Ini belum selesai  
Biasa saja, kamu tak apa

Bilas muka, gosok gigi, evaluasi (evaluasi)  
Tidur sejenak menemui esok pagi  
Walau pedih ku bersamamu kali ini <sup>B</sup>  
Ku masih ingin melihatmu esok hari

Bilas muka, gosok gigi, evaluasi (evaluasi)  
Tidur sejenak menemui esok pagi  
Walau pedih ku bersamamu kali ini  
Ku masih ingin melihatmu esok hari

