From IPA to AI:

A Linguist's Path from Academia to Tech

Alice Shen

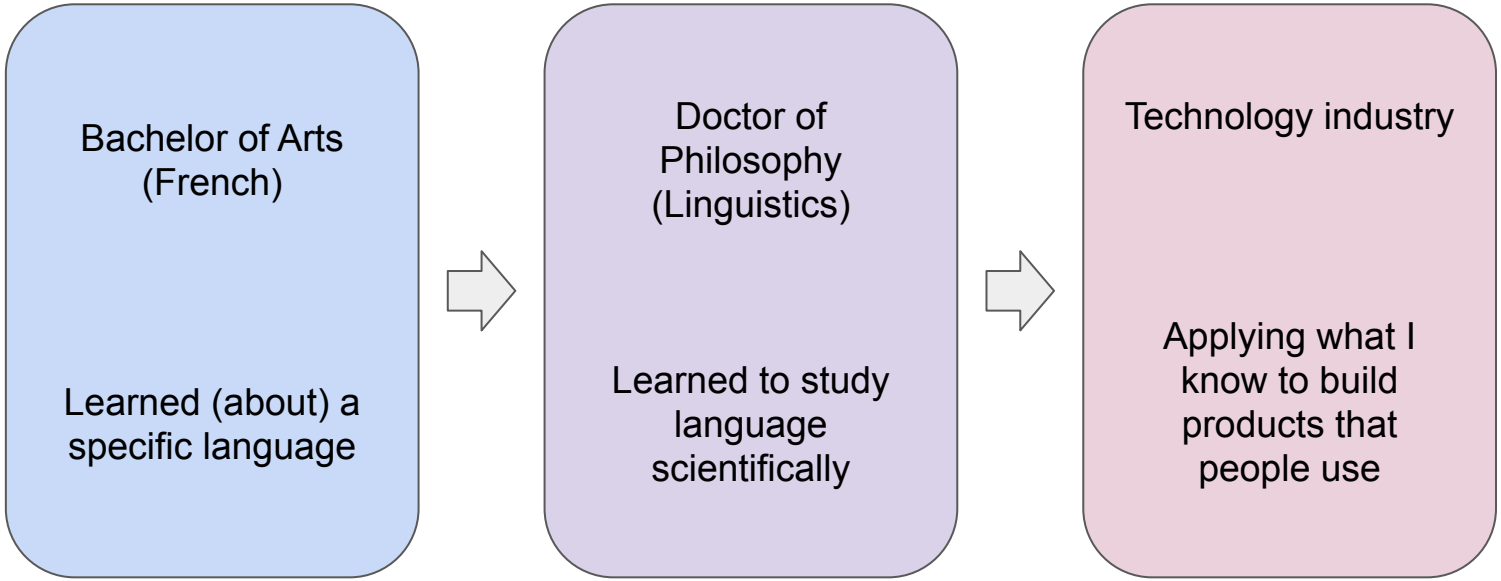
Linguistic Engineer @ Meta

Overview

- My path from academia into the tech industry
- Linguistic training in action
- Making the transition
- Zooming out on the industry

My path

From French verb conjugation to Voice AI

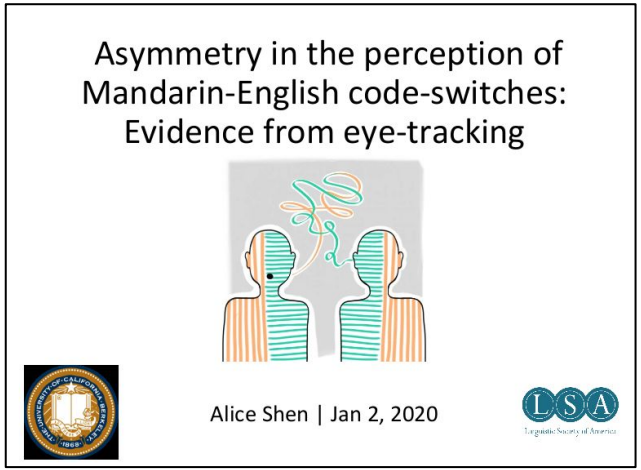


Undergraduate: Finding linguistics

Bachelor of Arts in French (2010 - 2014)

- I knew I wanted to study language
  - French language, literature, culture
- But I was also interested in other subjects
  - Math, philosophy, psychology, computer science, other languages
- I took a linguistics class one day and realized how interdisciplinary it is
  - Semantics = Math, logic, philosophy
  - Phonetics = Biology, physics
  - Sociolinguistics = Sociology, anthropology
  - Psycholinguistics = Psychology, neuroscience
- Linguistics combined everything I found interesting

Graduate: Diving into research



PhD in Linguistics (2014 - 2020)
Focus: Phonetics, bilingualism, code-switching

Research questions:
- How do bilinguals switch between languages in real-time?
- What acoustic cues signal a language switch?
- How does speech perception differ in bilinguals?

Skills that became a foundation:
- Data analysis: Processed speech data
- Statistics: Built models to test hypotheses (R Studio)
- Programming: Automated analysis pipelines (Python)
- Experimentation: Designed perception studies
- Communication: Taught courses, presented at conferences

Need to pivot:
- I loved the research process, but wanted my work to reach beyond academia

Academia to Industry Timeline

June 2020 — **Industry summer internship**
Linguist @ Facebook

August 2020 — **Received my PhD** 🎓🎉

One-year academic position
Visiting Professor @ Reed College

February 2021 — **Industry job search**

May 2021 — **Unemployment**

August 2021 — **Landed first industry job**
Analytical Linguist @ Grammarly

Industry internship: Testing the waters

Linguist @ Facebook (Summer 2020)

Challenge: categorize what people say into meaning categories that the voice assistant understands

What I did:

- Study how people talk to voice assistants
- Break down what they say into meaningful components
- Write instructions for data labeling
- Figure out what kind of utterances we were missing

Why it mattered: I realized I could apply my linguistics knowledge to building things that people use

This experience helped me land my first full-time role a year later.

First industry job: Learning the ropes

Analytical Linguist @ Grammarly (2021 - 2024)

Challenge: Used my problem-solving and research skills to build AI that improves people's writing in business settings

What I learned at Grammarly

Started: Figuring out industry in my first full-time tech job

Year 1:

- Figured out how to work with product managers, machine learning engineers, and data scientists
- Learned about the product development process
- Took on a medium-sized project from ideation to launch

Years 2-3:

- Came up with my own projects based on gaps I noticed
- Designed evaluation frameworks
- Led other linguists on larger projects
- Became the go-to person for certain features and requests

Ended: Confident in my skills, fluent in corporate work, and connected with a network that opened doors

Second industry job: Leveling up

Linguistic Engineer @ Meta (2024 - now)

Challenge: Evaluate quality, manage linguistic annotation projects, and deliver data-driven insights for speech recognition components in wearable AI assistants

Career move: Went from working on text AI to voice and multimodal AI, combining my strengths and interests and utilizing specialized knowledge from grad school

Linguistic training in action

Linguist @ Facebook

Half my job was understanding language (linguistics), half was organizing and analyzing data (tech skills).

- **SQL & Python (coding languages)** → Analyzed large amounts of text data to find patterns and problems
- **Linguistics knowledge (syntax & semantics)** → Determined why the model made mistakes and how to fix them
- **Vendor communication** → Coordinated with external teams to label and organize language data
- **Document writing** → Wrote clear guidelines for labeling aspects of language
- **Experimentation** → Tested different versions of AI models to see which worked better
- **Critical thinking** → Figured out what was causing problems and what to prioritize
- **Foreign language skills** → Worked on products that need to be localized to a specific region or expanded and internationalized

Analytical Linguist @ Grammarly

My linguistics training helped me understand what makes writing good, while my tech skills helped me teach that to AI at scale.

- **Built relationships across teams** → Collaborated with engineers and product managers
  - These connections later led to job referrals and interviews!
- **Scripting (writing code)** → Automated repetitive tasks like processing large datasets
- **Annotation management** → Coordinated teams, budgeted time and cost, set up annotation tasks
- **Guidelines development** → Wrote clear instructions for annotators to label or rewrite text
- **New feature research & development** → Researched how people write professionally, reviewed academic literature, and translated findings into new product features
- **Prompt engineering** → Iterated on crafting instructions to guide AI model outputs

(What is an Analytical Linguist?)

- Provides language expertise to teams building AI models
- Looks at language data and manages annotations
- Different from Computational Linguists who program linguistic rules

Case study – Confident Tone Rewrites project

- Researched linguistic markers of confidence vs. uncertainty
- Analyzed writing samples
- Created annotation guidelines: What makes writing sound confident?
- Managed teams labeling confident vs. unconfident examples
- Evaluated AI model outputs: Are the rewrites actually better?

Original	Rewrite
"I just wanted to check in..."	"I'm following up on..."
"I think we should consider..."	"We should consider..."
"Sorry, but could you..."	"Could you..."

Linguistic Engineer @ Meta

My phonetics and multilingual knowledge helps me understand where speech AI fails in different languages, while my data skills help me measure and fix it at scale.

- **Analyzed voice assistant performance** → Used phonetics to diagnose speech recognition errors across languages
- **Managed speech data projects** → Wrote SQL/Python scripts; coordinated annotation teams processing thousands of audio samples
- **Developed quality metrics** → Created evaluation frameworks and tracked improvement over time
- **Tested prompt strategies** → Experimented with different ways to get better AI outputs
- **Enabled internationalization** → Applied foreign language expertise to ensure products work globally

(What is a Linguistic Engineer?)

- Not a software engineer (don't write production code)
- Not just an annotator (do much more than label data)
- The bridge: Between linguistic expertise and machine learning systems
- Build datasets, design evaluation frameworks, drive technical projects

Making the transition

Transferable skills from academia

Technical	**Domain expertise**	**Professional**
● Data analysis ● Statistics ● Experiment design ● Scripting for data processing ● Creating data visualizations **Tools:** R, MatLab, Python/Jupyter, GitHub	● Identifying phonetic and phonological patterns in data ● Understanding multilingualism and code-switching ● Syntax/semantics knowledge ● Recognizing dialectal variation ● Multilingual competence	● Communication (teaching, presenting, writing) ● Collaboration (co-authoring) ● Project management (designing studies, meeting deadlines) ● Mentorship (training research assistants) ● Time management (juggling research and teaching) ● Navigating ambiguity ● Giving and receiving feedback (peer review, teaching evaluations) ● Self-direction (driving projects from research proposal to published paper)

Repurposing academic skills

Academia	**Industry**
Planning research studies ⇒	Planning data annotation projects
Teaching and training research apprentices ⇒	Onboarding annotators
Creating assignments and exams ⇒	Creating annotation guidelines
Literature review for research study ⇒	Literature review for product vision
Designing and conducting linguistics experiments ⇒	Designing and conducting annotation experiments + user experience surveys
Collaborating with co-authors ⇒	Collaborating with cross-functional partners
Writing research papers ⇒	Writing product feature quality reports

Valuable skills in industry

- Adaptability and resilience
  - Fast deadlines, difficult requests, and working toward a shifting launch schedule
  - Example: *Product vision pivots right before launch → you re-prioritize*
- Managing ambiguity
  - Vague requests, shifting new technology, constantly changing workflows (at a big company)
  - Example: *Software engineer requests a model evaluation → you define what to evaluate and how*
- Collaboration
  - Working with cross-functional partners while building trust and becoming an area expert
  - Example: *On a project team, you consistently meet deadlines and deliver high quality work*

CV vs. Resume

- Differences between my CV and resume:

Academic CV	**Industry resume**
● 7 pages	● 1 page
● Publications ● Presentations ● Education ● Skills ● Awards ● etc.	● Relevant work experience ● Relevant skills (incorporate keywords from the job posting) ● Education
● List everything I've done academically	● Highlight relevant experience and skills for specific job I'm applying to ● Omit prestigious academic awards and publications in famous journals

Reframing academic experience

For resumes:

- Focus on skills and impact, not just titles and duties
- Quantify when possible (# of students taught, hours of data analyzed, projects completed)
- Customize to each job - highlight what's relevant

Translate:

- Learn the STAR method (Situation → Task → Action → Result)
- Turn academic work into problem-solving stories:
  a. ❌ "I did research on code-switching"
  b. ✅ "I designed a study, recruited 50 participants, and created a dataset others now use"

Practice:

- Mock interviews with career services or peers
- Prepare 5-7 stories from your academic experience
- Get comfortable talking about your work to non-academics

Reality of breaking into tech

Observations:

- Referrals and connections can contribute to getting interviews
- Many job postings expect 1-2 years of experience
- Technical skills and behavioral skills are both important
- The market is competitive

My experience: Every job I've gotten involved knowing someone. I was referred for both my Facebook internship and current Meta role by a graduate department alumna. I wasn't referred to Grammarly, but a friend of a friend who worked there gave me a lot of interview tips.

This isn't to discourage you, but to help you be strategic. Build skills AND network.

Why linguists matter

Your strengths:

- **Different perspective:** Non-CS backgrounds are an asset, not a disadvantage
- **Language expertise:** Understanding *why* language works the way it does
- **Human-centered thinking:** Designing evaluation beyond just metrics
- **Edge case detection:** Spotting failures engineers miss
- **Cross-cultural awareness:** Building products for global users

Skills to add:

- Programming and data analysis
- Understanding of NLP/ML concepts (how AI systems work)
- Technical communication (bridging linguists and engineers)
- Adaptability and comfort with ambiguity
- Continuous learning mindset (tech evolves fast)

Your background in language and linguistics is the foundation. Add technical skills, and you become incredibly valuable to companies building AI.

Practical steps forward

Build connections:

- Network with alumni already in tech
- Maintain relationships with everyone you work with
- Seek informational interviews
- Create a LinkedIn profile and cold connect with anyone whose work you find interesting

Build a technical foundation:

- Learn to code (e.g., Python, SQL)
- Obtain credentials (courses, bootcamps, certificates)
- Competence matters more than perfection

Try to get your foot in the door:

- Internships
- Contract work (e.g., hourly annotation)
- Look into adjacent roles; can always try to pivot later

Start exploring industry now

- Read job postings to understand what roles exist
- Connect with alumni and other folks working in tech
- Attend career seminars and industry talks
- Learn about NLP, ML, and LLMs (concepts, not mastery)
- Practice data analysis and coding (Python, SQL) through coursework or projects
- Build your LinkedIn presence

Start early with 2-3 that feel doable! Exploration, learning, and building connections takes time.

Job titles to search

Linguistic Engineer • Computational Linguist • Analytical Linguist • NLP Engineer • Speech Scientist • Conversational Designer • Voice Designer • UX Researcher • Content Designer • Data Scientist • Prompt Engineer • AI Quality Evaluator • Language Data Specialist • Annotation Specialist • Technical Writer • Localization Specialist • Language Quality Analyst

Search tip: Use "linguist" OR "language" on job boards—titles vary widely across companies.

Zooming out on the industry

Current challenges

Global reach

- AI working in languages beyond English
- Most data is English; other languages under-resourced

Responsible AI

- AI perpetuates biases in training data
- Need inclusive design and evaluation

Shift to LLMs

- Work changed from rule-writing to evaluation and curation
- Bigger models = more need for linguistic judgment

Growing opportunities

Important areas for linguists:

- **Multilingual expansion** - Low-resource languages need expertise
- **Prompt engineering** - Optimizing human-AI interaction
- **Multimodal AI** - Combining text, speech, and vision
- **Voice & conversational AI** - AR/VR, assistants, new interfaces
- **Quality & evaluation** - Defining what "good" AI means

As AI becomes ubiquitous globally, companies need linguists who understand both language deeply AND can work with technology.

From IPA to AI: Worth It

When I started studying French, I had no idea I would end up working in tech.

Remember:

- Your "unconventional" path is actually your strength
- Linguistic training translates to tech more than you'd necessarily expect
- It's okay not to have it all figured out

My career motivation: Knowing my phonetics research helps people be understood across accents. Seeing my work enable communication in dozens of languages. Using what I learned in academia to build products that help people write better or provide accessibility.

Whether you take this path or another, remember: your deep knowledge of language is powerful. The world needs people who understand how humans communicate—especially as we build the AI systems of the future.

My contact info

LinkedIn: https://www.linkedin.com/in/**azshen**/

Website: https://sites.google.com/view/aliceshen/home

Email: az.shen93[at]gmail[dot]com