Managing an Interdisciplinary Ph.D.

Tips and Tricks

William N. Havard

ALPS 2023 Winter School - 20/01/2023

william.havard@ens.psl.eu
william-n-havard.github.io

About Me

- A-Level in Maths, Physics, Chemistry, and Biology
- B.A. in German Language, Literature, and Linguistics
- M.Sc. in Speech and Language Processing
- Ph.D. in Computational Linguistics
  - “Lexical emergence from context : exploring unsupervised learning approaches on large multimodal language corpora” (05/2021)
  - Studied **child language acquisition** (spec. lexical acquisition) using **deep learning approaches**
  - My two fields of research:
    - **(Psycho)Linguistics (L)**
    - **Computer Science (CS)** and more specifically **Speech Processing**

#1 Catch Up!

- Depending on your background, you might lack knowledge in linguistics or CS

- **Catch up!**
  - Background in **Linguistics (L)**
    - Computer programming (Python, etc.)
    - Maths basics (algebra, calculus, stats, etc.)
    - Machine Learning (neural networks, etc.)
  - Background in **Computer Science (CS)**
    - Phonetics, Phonology, Morphology, Syntax, Semantics, etc.
    - Psycholinguistics, sociolinguistics, etc.
    - e.g. *Linguistic Fundamentals for Natural Language Processing* (E. M. Bender)
  - For both: **know the basics of your field (GMM, HMM, etc.)**
    - e.g. *Speech and Language Processing* (Jurafsky and Martin)

➢ Otherwise you might end up using some terms incorrectly
  - e.g. *"On the use/misuse of the term 'phoneme'"* (Moore & Skidmore)

#2 Be Selective! (1/2)

- Doing interdisciplinary research is **challenging** because you have two (fairly) distant fields you are trying to bring together

- Which supposes you know about both…but **you can't read everything and you can't understand everything in one field, let alone two…**

- **Be Selective!**
  - **Read publications/books relevant to YOUR topic**
  - e.g. "language acquisition"
    - → psycholinguistics (*maybe* reading theoretical paper on syntax, despite being very interesting, is not very useful)
    - → computational models of language acquisition (so *maybe* knowing the theory behind neural networks in and out is not useful)

#2 Be Selective! (2/2)

- Stay **focused**!
- Know **what depth of details is necessary & required** for your work

- "A good Ph.D. is a finished Ph.D."
- You'll have **plenty of time after your Ph.D.** to read and learn more about the intricate details you had to skim over during you Ph.D.

#3 Try to Keep It Balanced

- Try to **keep your reading balanced between both fields**
  - It's **OK if it's not 50/50**, we all have a favourite subject
  - At some point you might end up doing 65% CS reading and 35% L, and later on the ratios will reverse

- Try **to read regularly in both fields**
  - e.g. let's say **for every 2 CS paper you read 1 L paper**, or vice versa
  - **Except at the beginning** when you have a lot of catching up to do
  - You have fruitful ideas when you read regularly and sometimes an article changes your perspective on your current experiments

- If you don't manage both at the same time, **work in shorter chunks**
  - e.g. 2 weeks CS, 2 weeks L

- Balancing your reading **helps you stay up-to-date in both fields**…

#4 Stay Up-To-Date (1/2)

- You can't know everything on both subjects … but **you shouldn't miss out on the big trends**
  - **Technology:** Transformers, BERT, Whisper, wav2vec-U, CPC, Adapters, …
  - **Methodology:** BERTology, Fine tuning, Transfer learning, low-resourced languages, …

- As said before … **you don't have to know these new technologies/methodologies inside out**.

- **Enough to understand the main concepts** and papers that use them and how you could apply them to you work

#4 Stay Up-To-Date (2/2)

- **How?**
  - ArXiv mailing-list
  - Attend conferences and spot the big trends
  - Talk to your lab mates
  - Talk to your supervisor(s)
    - They have a broader view and understanding of your field than you
    - They are able to spot what might catch on what won't
  - Read papers that were awarded best paper/presentation/whatever awards
  - Become a reviewer for a top-tier conference!

- **Do the same for your other field** (luckily, it does not evolve that fast in L)

- Sometimes news technologies/models can be daunting as they require a lot of prior knowledge
  - **Use "predigested" knowledge** (blogs, videos) that help you get the big picture without too many details
  - Easier to grasp the big intuitions and go into the details afterwards (if needed) than the other way round

#5 Know Your Audience

- Because your doing an interdisciplinary Ph.D. your work will bring contributions to two differents fields

- The audience you are talking to **might not be familiar with the concepts of your other discipline**
  - e.g. NN for linguists; psycholinguistics for CS people

- **Orient your talk**/presentation/poster/paper according to the audience at the venue
  - Focus on computational contributions in a CS venue
  - Focus on linguistic contributions in a L paper
  - **Show the potential of the other field to the venue**

#6 Be Creative

- **2 disciplines = 2 methodologies**
- Take methodologies from one domain **apply it to the other**
- In my case, everytime I read a (psycho)linguistics paper I think of **how the methodologies described in the paper could be applied to analyse NNs**
  - usually requires a fair amount of adaptations…
  - **fruitful and novel results**
  - you already have a human baseline you can “compare” to

#7 Listen to Feedback

- **Listen to Feedback!**
- Particularly from researcher in the **field you are less comfortable with**
  - Make sure your research questions **make sense to both fields**
  - **Are they getting the point of your research?**
  - **If not…**
    - Maybe **you didn't make your research questions understandable enough to them**
    - Maybe **your research has drifted too much toward your other field** and your research questions don't "resonate" with both fields anymore

QUESTIONS

Question: Literature Review

- When should I stop reading generalist documents?
  - It is normal read generalist documents (i.e. not directly relevant to you research questions) **at the beginning**, especially if you have a lot of catching up to do
  - Your reading should gradually narrow down to more focused articles/book chapters pertaining to the specific research questions you aim to answer
  - If not
    - **Your research question is too broad/general**
      - you either try to encompass too many aspects of your subject at once
      - or you fail to see the individual aspects of your subject
    - **Talk to your supervisors!**
      - They can help you narrow down your research questions and help you put your ideas into words

Question: My Supervisors Don't Agree

- What should I do if my supervisors don't agree?
  - **Make sure you're all on the same page**
    - Meet with your two supervisors at the same time
    - **Make sure they understand what the other is talking about**
      - You should serve as a bridge between them and their respective area of expertise
      - "Translate" what the other is saying in understandable terms
  - If they still don't agree
    - They might have a **different research agenda**
    - **Try to find a compromise**

Question: Becoming a Reviewer

- How can I become a reviewer?
  - Complain on Twitter
  - Contact the ACs
  - **Become a sub-reviewer**
    - Main-reviewers can “delegate” some of their review duties to sub-reviewers
    - Ask your supervisors to give you one of the paper they were assigned to review
    - They usually have a lot on their plate and **will be happy to delegate**
    - **They can give you feedback on your review** (“you missed that point here”, “you were too harsh here”, inadequacy between your review and scores, etc.)

Question: Catching Up

- How do I catch up?
  - Use a lot of **“predigested” knowledge at first**
    - **Blog posts / Tutorials**
      - medium.com / towardsdatascience.com /author’s own personal blog
    - **Videos**
      - lots of free and quality content en YouTube
    - MOOCs
      - but you generally have to pay, which I do not recommend given the wealth of freely accessible resources
  - Choose an introductory course at your university (Python101, Ling101,...)
    - Attend master’s level lectures (required in some universities)
  - **Identify your weaknesses as soon as possible**

Question: Keeping Your Research Balanced After Your Ph.D.

- How do I keep my research balanced after my Ph.D.?
  - Find a company that welcomes your interdisciplinarity (if you want to work in industry)
  - **Find a postdoc position in the field you have less addressed during your Ph.D.**
    - In my case, still doing interdisciplinary research
    - But with a greater focus on linguistics and psycholinguistics

Question: Reading Groups?

- Do you recommend attending reading groups to catch up?
  - In any case, I think it's a good idea to attend reading groups if you manage to make room for it
  - **It will force you to read regularly**
  - Maybe attending a reading group with a majority of senior researchers might be hard at the very beginning
  - Create your own reading group with fellow Ph.D. students
    - **Reduces your reading burden**
    - **Trains you** to present your future papers
    - Everyone benefits from the other's readings