

Reflective Research Report: Improving Chemistry Lab Script Accessibility for ESL Students

Introduction: The Importance of the Project

Anyone who has studied a science knows how daunting the feeling of walking into a lab can be - the foreign equipment, strange machinery, complex instructions and dangerous chemicals. Now imagine walking into the same lab, but this time the equipment, the machinery, the instructions and chemicals are all in a language that is not your native one. Immediately an already daunting experience would become even scarier. This leads us to the goal of my research to improve the accessibility of chemistry lab scripts for students who have English as a second language (ESL). To create a lab environment that is accessible and comfortable for all.

Chemistry is a language-heavy subject, filled with specialized terms, complex instructions, and technical jargon. While this language can be difficult for native English speakers, it presents an even greater challenge for ESL students. At the University of Leeds, where 10-15% of the undergraduate chemistry cohort consists of international students, it is crucial that these students are able to fully engage with their lab content. Laboratory scripts that are difficult to understand can lead to confusion, frustration, and yes poor academic performance- but more importantly low self-confidence.

Clear and accessible instructions are vital for ensuring that all students, regardless of their language proficiency, can perform experiments safely and correctly. Therefore, my research aimed to analyze the language used in chemistry lab scripts and develop strategies to make them more accessible for ESL students, improving both learning outcomes and safety in the lab.

Initial Meetings and Research Planning

Before my project, I thought it beneficial to contact universities nationwide, in order to ensure my research explored the overall picture of how accessible our UK chemistry lab curriculum is for ESL students. I contacted over 60 universities, explaining the purpose of my research and requesting access to their Year 1 undergraduate chemistry lab manuals. While I only received responses from 10 universities, this gave me a comprehensive and diverse view of the issue. This stage taught me valuable project management skills (that I had previously never used) including how to maintain professional correspondence, create effective information sheets, and manage the logistics of gathering external data.

Through this experience, I learned about the challenges of engaging external contributors in academic research. The low response rate, while frustrating, encouraged me to develop resilience and a proactive approach in following up with potential contributors. Reflecting on this, I realized that persistence is crucial in research, especially when relying on external sources. This lesson will undoubtedly benefit my future research endeavors, emphasizing the importance of adaptability and the ability to engage effectively with various stakeholders.

On day one of my project, I met with my project supervisor, Dr Keith Livingstone, and Dr. Jenny Eyley, another member of the research team. I was nervous about the project, expecting that they would outline specific tasks for me to complete. However, both Keith and Jenny gave me complete freedom in deciding the direction of my research. This was both exciting and daunting. I'm someone who is used to having a clear set of goals, so being

given such autonomy was initially intimidating. For the first time, I had to rely on my own critical thinking to shape the research, which pushed me out of my comfort zone and allowed me to develop leadership skills that I had not previously used.

Keith and Jenny introduced me to a software called **SketchEngine**, a tool used for linguistic analysis. SketchEngine works by feeding lots of texts (in this case lab scripts) into SketchEngine, to create a large database of language called a corpus. SketchEngine then compares the corpus to the English language- identifying the most frequently used words in the lab scripts, and generating a list of words used most frequently in the lab scripts, compared to not so frequently in everyday English Language.

At first, I found the software overwhelming. I had never worked with such a tool before, and the abundance of unfamiliar terms and functions left me feeling slightly disappointed. I had hoped to dive straight into the research, but I quickly realized the importance of taking the time to understand the software and build a solid foundation. Over the next few days, I watched tutorials, learned key terms, and met with Jenny to discuss the details of the software and my research ideas. This process of preparation was crucial and helped me appreciate the value of being thorough and methodical in research.

I also created an initial master list of research tasks, which helped me organize my thoughts and plan my next steps. Keith suggested that I prioritize the tasks and look for ways to complete them in tandem to maximize my efficiency. This advice helped me see a clearer path forward and allowed me to break the project into manageable chunks.

Summary

- categorise words-vocab,procedure,equipment etc
- create corpus and sub corpii
- reading age of each lab script
- reading age of each uni-come up with statistics
- look at noun phrases
- look at percentage of words for comprehension
- look at chemistry dictionary-what words dont I know
- compare chem word dictionary with frequency in lab scripts
- repeat with IELTS
- create a metric with this
- what words fall through the cracks, how many words can they now understand in the lab script
- look at creating a vocab list
- look at average number of words per lab script
- summary document
- areas of good practice
- look at layout-just pure vibes and highlighter
- look at start of QMU
- upload all to onedrive
- look at literature to provide context for presentation and ideas for me

One key piece of advice from Jenny was to always keep my "why" in mind. When conducting research, it can be easy to get lost in the details and lose sight of the overall goal. By constantly asking myself what conclusions or outputs I could produce from my work, I stayed focused on the larger purpose of my project: improving lab script accessibility for ESL students.

Early Progress: Literature Review and Readability Testing

At this point I was told I was to present a project proposal to the Chemistry Department's pedagogical team at the end of the week , and therefore I had to had to have some preliminary research conducted.

Therefore, in the first week of my project, I focused on two main tasks: conducting a **literature review** and testing the **readability** of our lab scripts. I wanted to gather evidence to demonstrate the importance of my project to the Chemistry Department's teaching team. This was particularly important, as I would need to convince experienced chemists—who may take linguistic understanding for granted—that improving lab script accessibility was a necessary task.

For the literature review, I focused on research that highlighted the challenges ESL students face in academic settings, particularly in STEM disciplines. I found several studies showing that language barriers can significantly hinder students' ability to understand technical instructions, even when they have the required subject knowledge. One particularly relevant finding was that for students to understand 60% of a text, they need to know at least 94% of the words. Most interestingly, I found that the UK home office suggests that to be easiest to digest, texts should have a reading age of 9! With this context in mind, I decided my first task would be to test the readability of our chemistry lab scripts.

I used the **Flesch-Kincaid readability scale** – a scale that gives a text a score from 1-100 with 100 being the easiest to read. I evaluated the readability of lab scripts from the University of Leeds. This scale assigns a score based on sentence length, paragraph structure, and word complexity, with a higher score indicating easier readability. For international students, Leeds requires a minimum score on an English Language Proficiency test (IELTS) – a score of 6.5, which roughly equates to a Flesch-Kincaid score of **65**. However, after removing technical chemical names and equations (to ensure I only focussed on parts of our lab scripts that could be changed), I found that our lab scripts had an average score of **55**, well above the level of English comprehension required for entry to the course. This result clearly demonstrated the need for improvements, as students were expected to work with materials that were harder to read than their language skills allowed.

Presenting these initial findings to the Chemistry Department's teaching team was a highlight of my first week. It was a nerve-wracking experience, as I was speaking in front of experienced chemists, but their constructive feedback on how to improve the structure and visual aids in the lab scripts was invaluable. The presentation reinforced my commitment to making lab scripts more accessible and provided me with useful insights into how to move forward with the project.

Linguistic Analysis: Identifying Vocabulary Barriers

Over the next two weeks, I conducted a detailed **linguistic analysis** of the lab scripts using SketchEngine. My goal was to identify vocabulary that could be particularly challenging for ESL students. To do this, I first generated a list of the most frequently used words across the lab scripts, filtering out words that were likely familiar to ESL students by cross-referencing them with the **IELTS vocabulary list**. This process allowed me to isolate specialized terms that were likely to be unfamiliar and challenging. I then used Sketch Engine's Keywords function to compare this new list with a reference corpus of the everyday English language, to produce a list of words that are used most frequently in our lab scripts, compared to not frequently at all in the English Language. This generated a list of words most likely to be unfamiliar to ESL students.

My next task was to create a glossary of these terms by categorizing the top 25% of identified words into various themes:

- **Chemicals and Reagents**
- **Scientific Concepts and Names**

- Physical and Chemical Properties
- Analysis Terminology
- Lab Jargon
- Lab Equipment
- Variables, Constants, and Measurement Units
- Techniques and Processes
- Lab Safety
- Academic Instruction

Through this categorization, I gained a better understanding of the areas where ESL students might struggle. For instance, there were many complex words being used to describe simple techniques and processes, or academic instruction (the word 'transcribe' being identified as unfamiliar and yet used a lot). These findings emphasized the importance of simplifying and clarifying lab instructions to avoid misunderstandings. This led me to go through Leeds University lab scripts by hand, changing this unnecessarily complex vocab.

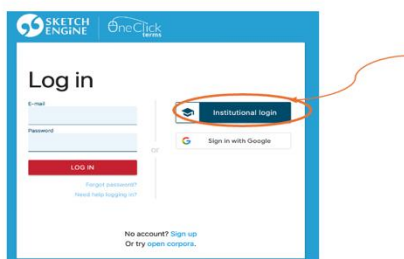
This list was extremely useful to me, and therefore I knew it would also be for staff members at other unis who had expressed a keen interest in increasing the accessibility of their own scripts. However, I found it difficult to make the list, so I decided to come up with simple instructions I could send out to staff members. I ensured to focus on the adaptability of the list – changing it's reference corpus, categories or uses could benefit staff in different ways.

Instructions for Generating a Glossary of Unfamiliar Lab Script Vocabulary for ESL Students

The following steps outline how to generate a glossary of vocabulary that may be unfamiliar to ESL students using **Sketch Engine** and **Excel**.

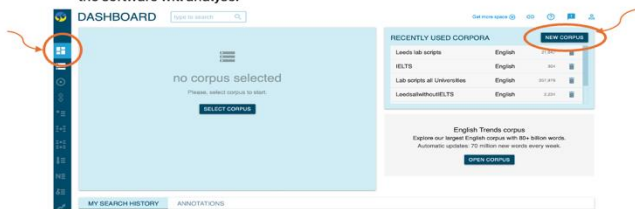
Logging into Sketch Engine

Use your university ID to login and gain full access to the platform



1. Generate a Corpus of Lab Scripts

- Upload all your lab scripts into Sketch Engine to create a 'focus corpus.' A corpus is a database of language, and the focus corpus is the collection of language that the software will analyse.



2. Create a Wordlist of Frequent Terms

- Use Sketch Engine's **Wordlist** function to generate a list of the most frequently used words in the lab scripts.
- Ensure **'word'** is selected-this allows us to see multiple forms of the same root word (e.g., "precipitate" vs. "precipitates"), which may have different meanings in lab settings.



- Export the wordlist to **Excel** for further analysis.

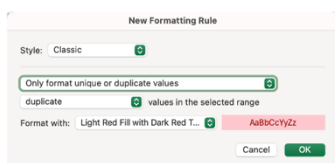


3. Create your Cross-Reference List

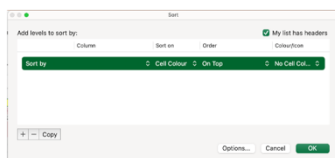
- In this case we used the IELTS vocab list-a list of words students should know for the university-required level of English language proficiency. Although we don't expect students to know every word on this list, it does allow us to only focus on words that will be especially unfamiliar to students
 - The link to the wordlist we used - <https://www.ielts-up.com/academic-wordlist.pdf>
 - You can use any cross-reference list here, depending on how you wish to analyse your lab scripts. If you're trying to design your text to a different standard, this is the divergence point
- Export this wordlist to Excel and place it in the column next to the lab script wordlist.

4. Compare Lists

- You will see excel shows you a frequency column for each word, whilst this is interesting and could be of use in future, it is not necessary for this process and can be deleted.
- In Excel, select both columns and use **Conditional Formatting** to highlight duplicates between the two lists. These are words that appear in both the lab scripts and the IELTS corpus.



- Apply a **Custom Sort** to group the highlighted (duplicate) words (in the Lab Script column) together.



- Here is a link to some useful excel guidance on Custom Sort and Condition Formatting

- Sort- <https://support.microsoft.com/en-gb/office/sort-data-in-a-range-or-table-62d0b95d-2a90-4610-a6ae-2e545c4a4654#:~:text=In%20the%20Custom%20Sort%20dialog,font%20color%2C%20or%20cell%20icon.>
- Conditional Formatting- <https://support.microsoft.com/en-gb/office/highlight-patterns-and-trends-with-conditional-formatting-eea152f5-2a7d-4c1a-a2da-c5f893adb621#:~:text=formatting%20to%20text,Select%20the%20range%20of%20cells%2C%20the%20table%2C%20or%20the%20whole,highlight%2C%20and%20then%20click%20OK.>

5. Remove Known Vocabulary

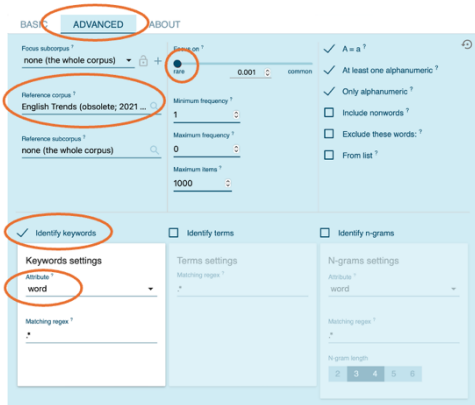
- Delete the duplicated words, narrowing down the list to terms that are least likely to be familiar to ESL students.
 - Note: This step helps reduce the number of words analysed, though we cannot guarantee that ESL students know all the IELTS vocabulary.

6. Create a New Corpus from Unfamiliar Words

- Use the list of non-duplicated words (the remaining lab script-specific words) to generate a new **corpus** in Sketch Engine.
 - Excel is not a suitable file format for Sketch Engine, so copy and pasting into a word document or exporting to PDF are both suitable methods.

7. Identify Keywords Using a Reference Corpus

- Use Sketch Engine's **Keywords** function, under the 'advanced' tab, setting the rarity threshold to **0.001**, to focus on terms that are uncommon in general English.
- Choose a recent reference corpus, with a large number of words like **English Trends (obsolete; 2021-March 2024)**. This ensures you're comparing your lab script vocabulary against contemporary, accurate English usage.
 - Deselect** terms and N-grams (you can run this analysis later if needed).
 - Select **Word, not Lemma**. This means you'll differentiate between different forms of words (e.g., "precipitate" vs. "precipitates"), which may have different meanings in lab settings.



8. Generate and Export the List of Keywords

- The list will contain the most common words in your lab scripts that are least common in general English.
- Export this new list to **Excel**.

You have now completed the necessary steps to generate a list of words most likely to be unfamiliar for students. From here forward it is up to you to utilise this, it could be that you notice some are essential to your scripts, however others can easily be replaced with a simpler word, or you may look at creating resources or learning activities for students, to help familiarise them with the words. We thought it useful to categorise the words for future use:

9. Categorize the Top Keywords

- Focus on the **top 25%** of words from this list (these are the most unfamiliar to students). Categorise the words into the following categories (and anymore you see fit) using Cell Fill function

Key	
Yellow	Chemical
Orange	Equipment
Green	Property
Blue	Technique
Purple	Other
Light Blue	Lab-specific jargon
Dark Blue	Processes
Light Green	Reaction Type
Light Purple	Lab safety and Procedures
Light Orange	Time words
Light Yellow	Chemical Agents/States
Light Blue	Instrument Units and Variables and G
Light Green	Analysis
Light Purple	Scientific Laws and concepts

- Use **Custom Sort** to organize the list by cell colour, so you can easily see all categories.

10. Create a Glossary

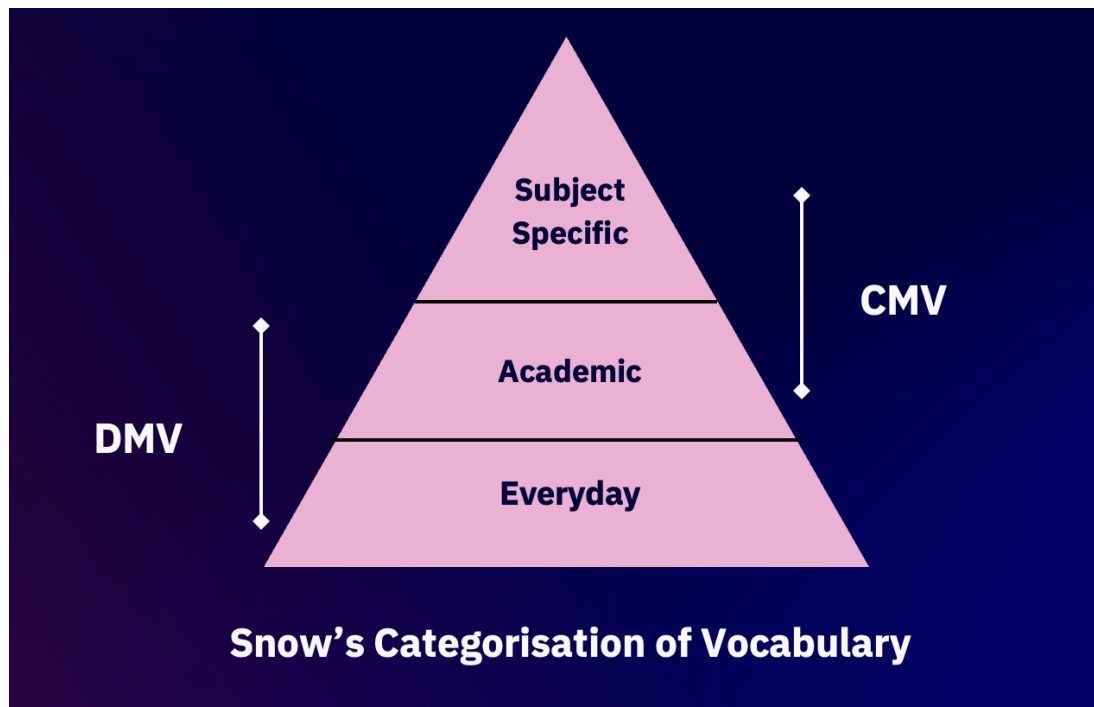
- Export each categorized section into a **glossary**.
- This glossary could then include definitions or explanations of the terms in simpler language, or even provide translations or visual aids if necessary.

My supervisor kindly tested out my instructions and initially found them difficult to follow, as because I had been doing this project for 3 weeks, certain things seemed obvious to me that were not for other people. This taught me the importance of placing myself in the audience's shoes and ensuring I communicated my ideas

clearly and effectively.

Categorisation of Vocabulary

In addition to categorizing vocabulary, I came across something very interesting in my initial literature review called Snow's categorisation of vocabulary.



It splits words into 3 tiers- everyday language, academic terms and subject specific terms.

Tier I includes words used in everyday life by the majority of the English-speaking population, these are words that students generally learn at an early age or in English language courses .

Tier II vocabulary is the vocabulary used in academic culture, across the disciplines.

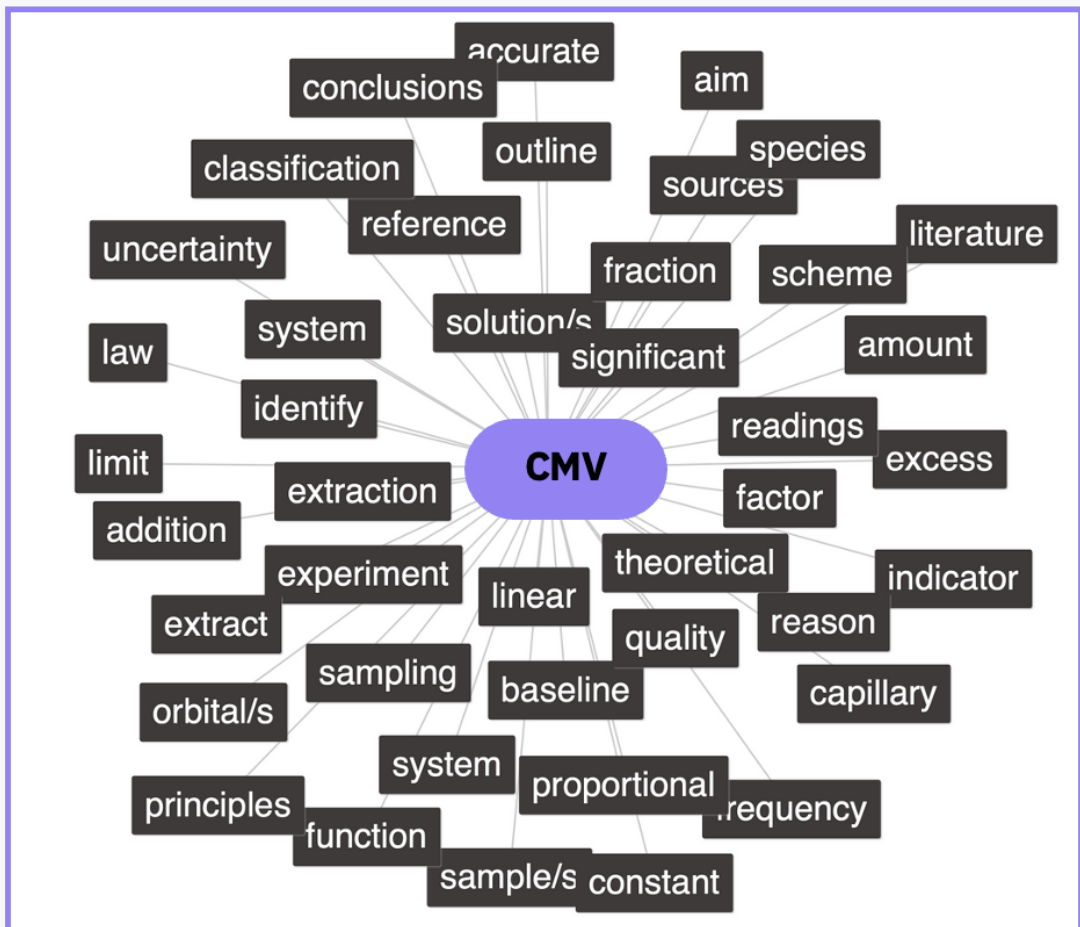
Usually these words are not taught explicitly in classrooms, yet students start learning them through being exposed to an educational environment. Tier III are discipline-specific words - scientific words are considered Tier III words.

Confusingly, some Tier III, scientific words have multiple meanings. Acid, sugar, and mole have technical Tier III meanings in chemistry but different meanings when used in a Tier I everyday setting. These are called **Dual Meaning Vocabulary (DMV)**.

Even more confusingly, some words have academic meanings (Tier II) and technical meanings (Tier III). Often, scientists use Tier II words to create specific terms for Tier III concepts (e.g., indicator). While the definitions are similar, discussing these words in an academic setting usually implies a different meaning than when discussing them in a technical setting like in the lab. This is called **Cross Meaning Vocabulary (CMV)**.

DMV and CMV are particularly tricky for ESL students because they are familiar, yet their meaning in the lab context can be completely different. So I set out to identify them across all ten universities' lab scripts. I used sketch engine to create a list of the most frequently used words across all the scripts, then I worked through manually, identifying all CMV and DMV.


Here are the results:



Identifying and categorizing these types of vocabulary deepened my understanding of the language barriers ESL students face in technical subjects. The work also highlighted the importance of contextual understanding—knowing a word isn't enough; students need to understand its specific meaning in the context of chemistry to successfully complete lab tasks. I realised that I myself had never heard of this issue and yet it was such a prevalent barrier, so I set out to create an infographic to educate other staff members at Leeds and other unis on the issue.

Breaking Down Barriers


Blue and Green Minimalist Modern Business
Infographic Poster - Poster



Supporting ESL Students in the Lab


THE CHALLENGE

For ESL students, navigating lab scripts can be a real challenge. Why? Many words in these scripts have **multiple meanings**, making it easy to misinterpret critical instructions. Our research focuses on addressing these vocabulary hurdles to **improve understanding and accessibility** for all students.



WHY THIS MATTERS

The **shift in meaning** between **everyday**, **academic**, and **technical** contexts can create significant **confusion** for ESL students. Words they recognize in one setting take on entirely different meanings in another, making it harder to follow lab instructions and understand core concepts. This issue is not exclusive to ESL students, this vocab can be confusing for **all new** lab students.



DMV

Some words seem familiar, but their meanings can shift dramatically in a lab setting. These are known as **Dual Meaning Vocabulary (DMV) words**. For example:

Acid

- Everyday: Found in sour foods, like lemons.
- In the Lab: A substance that donates protons in a chemical reaction.

Mole

- Everyday: A small animal or a spot on your skin.
- In the Lab: A unit to measure the amount of a chemical substance.

WHAT'S THE ISSUE?

Other words have academic meanings but take on technical definitions in the lab. These are **Cross Meaning Vocabulary (CMV) words**:

Indicator

- Academic: A way to solve a problem.
- In the Lab: A mixture of substances in liquid form.

Experiment

- Academic: A trial or test to explore a theory.
- In the Lab: A detailed procedure to confirm scientific data.

CMV


HOW CAN WE ADDRESS THIS?

We've identified a few key strategies to make lab scripts clearer and more accessible for ESL students:

- **Create Easy-to-Use Glossaries:** Develop glossaries that explain both the academic and technical meanings of DMV and CMV words, complete with examples from lab scripts.
- **Provide Context Clues:** Include brief clarifications within lab scripts to flag words with multiple meanings, helping students understand the correct usage.
- **Develop Phrase Banks:** Analyze common word patterns to create reference materials that show how these tricky words are used in real lab situations, making learning more intuitive.

THE IMPACT

By focusing on DMV and CMV, we can reduce the effect of language barriers and help all students—especially ESL learners—**better engage** with lab materials. Through clearer vocabulary, we open the door to **greater understanding, academic success**, and a lab environment that is **comfortable** for all students.



Simplifying Lab Script Language: Standardizing Terminology

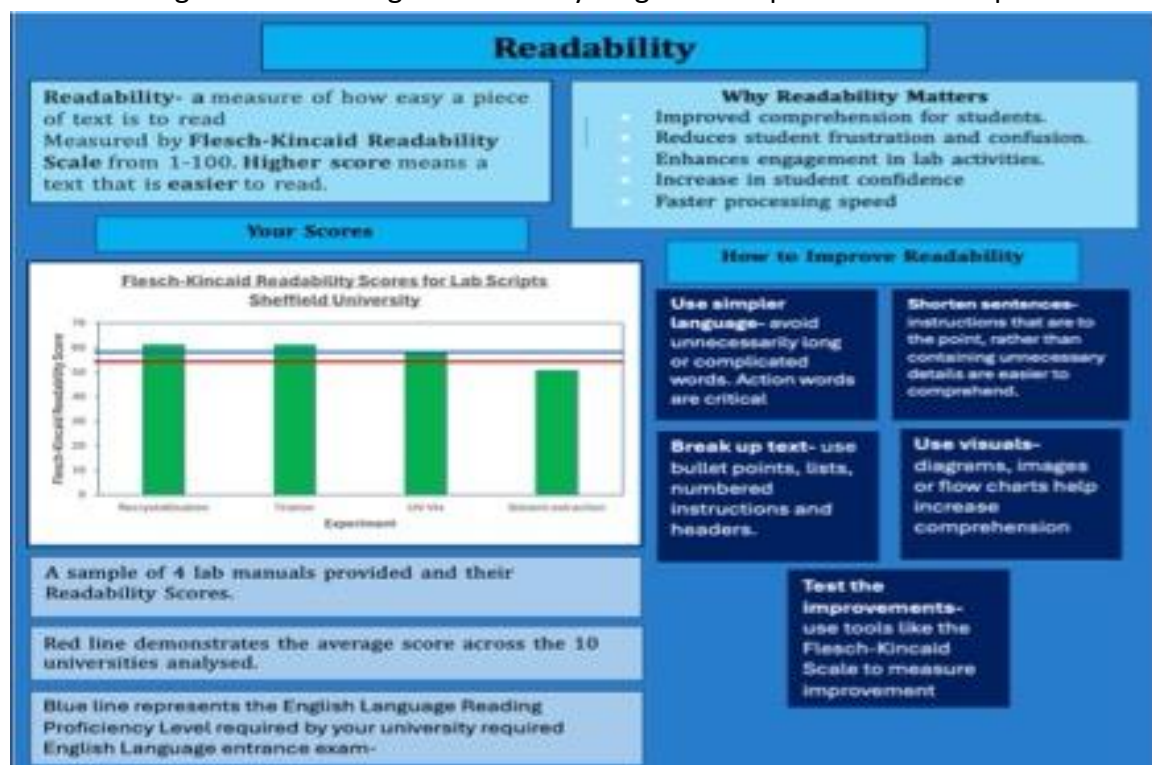
One of the most crucial phases of my project was **simplifying the language** used in lab scripts and **standardizing terminology**. From my linguistic analysis, I found that many lab scripts used multiple terms for the same technique, such as referring to "vacuum filtration" as both "Buchner filtration" and "suction filtration" This inconsistency adds to the cognitive load for ESL students, making it harder for them to follow instructions. Luckily, my chemistry department were creating a basic skills manual for year one students, so I ensured that under each skill covered, I also added any other ways it could be referred to.

I also focused on simplifying sentence structures. The lab scripts often contained dense, complex sentences, which could overwhelm students. By rewriting sentences in the **imperative voice** (e.g., "Reheat the filtrate" instead of "The filtrate is reheated") and breaking up dense paragraphs into bullet points or numbered lists, I made the instructions more straightforward. These changes aimed to reduce ambiguity and make the steps easier to follow for all students.

Through this process, I reflected on the **power of clear communication**. Simplifying complex language does not mean dumbing it down; rather, it makes the information more accessible to a broader audience, ensuring that all students can engage with the material effectively.

Readability

Finally I analysed the readability of a selection of lab scripts from each university(ensuring my collection contained the same experiments across the unis and covered a wide range of labs). Increasing the readability of a script is a surefire way to increase student comprehension and so I decided to research how universities can do this. Luckily, it has nothing to do with the content of the script, but more the layout. Things like breaking up large bodies of text by adding numbers or lists, using images or flowcharts, reducing sentence length and removing unnecessarily long and complex words all helped readability. I



knew this would be vital information for staff, so I created a dashboard for each university-

covering their readability scores, how they compared to their English Language Proficiency required for admission and how they compared to the average across all ten universities. I also added ways to improve their scores.

The Impact and Future of the Project

Overall, this research project has had a significant impact on both my personal and professional development. By conducting a linguistic analysis of chemistry lab scripts and developing strategies to make them more accessible, I have contributed to improving the educational experience for ESL students at the University of Leeds. This project has also highlighted the importance of ongoing efforts to make scientific education more inclusive, ensuring that all students, regardless of their language background, have equal opportunities to succeed.

I partook in a very exciting meeting with the university's language department in my final week of research, in which I found out that the wordlists and glossary's I created will be used in ESL specific learning workshops and used to create resources and activities. Furthermore, a master's student is starting a project using my findings to create learning resources and activities for ESL students in Chemistry, I will be entrusted with educating him on my findings. I also plan on attending national chemistry teaching conferences such as the CLEAR lab symposium and ViCEPHEC to present my findings, I hope to engage with a broader audience and contribute to the growing conversation around accessibility in STEM education.

Through this project, I have learned that small changes in how we communicate complex information can have a profound impact on students' learning experiences. By making lab scripts clearer and more accessible, we can help ensure that all students, regardless of their language proficiency, have the tools they need to succeed in the challenging world of chemistry.

Reflection: Professional Growth and Leadership

As the project progressed, I found myself developing **leadership and project management skills** that I had not previously used. Being given full responsibility over the direction of the project was initially daunting, but as I gained confidence, I realized the value of being able to shape the research according to my own vision. This autonomy allowed me to think critically about the work I was doing, constantly refining my approach based on feedback and new insights.

I also learned the importance of **collaboration** and **communication**. Presenting my findings to the Chemistry Department's teaching team and engaging in discussions with faculty members and other researchers gave me valuable feedback and helped me build strong professional relationships. I became more comfortable presenting my ideas and more adept at tailoring my communication style to different audiences, whether I was speaking to experienced chemists or the language department.

Additionally, this project taught me the importance of **reflection** in research. By constantly revisiting my goals and keeping my "why" in mind, I was able to stay focused and ensure that my work was driven by a clear purpose: improving accessibility for ESL students.