

Hypothesis and data collection

A vital part of your project will be your research question and/or your hypothesis/hypotheses (plural). This will be a question or statement near the beginning of your investigation that will frame your project and help to focus your research. They will also give your project a sense of structure and show you're working with a **methodology** - just what the teacher marking your investigation wants to see!

Formulating a research question

A research question will help to demonstrate precisely what the focus of your investigation is. It's far too vague for you just to say you're doing an investigation on language and gender for example. Often it is accompanied by a hunch about what the possible answers might be (see hypothesis below). However, it is important to ask yourself open question(s) as your initial thoughts can be wrong! For example, in much gender research (say for example in gender and spoken language) the question posed by the researchers assumes that there is a difference between the speech styles of men and women.

Question:

Can you think of a reason why it is not good practice to assume that there is difference in male and female speech? (How might posing a research question like that affect your research?)

What is a hypothesis?

It's likely that once you've chosen your topic and considered a research question, you'll already have a good idea of what you'll find (probably due to your study of A-level English language, or because it's a topic with which you're familiar). This is where you can give a hypothesis or hypotheses - educated, linguistic guesses as to what you expect to find. For example, if you're analysing how teachers maintain authority in the classroom, you could reasonably expect a high level of imperatives, or a use of low frequency lexis.

When structuring a hypothesis or research question, it is important to consider the following:

- Do you have an expectation as to what you might find with your topic/chosen data? If yes, your hypothesis can be a more precise expected statement of fact: 'African-American rappers will employ AAVE language features in their lyrics to construct their identity'.
- If you have no idea what to expect in your data and chosen topic, you will find it much more useful to employ a research question: 'To test whether African-American rappers employ AAVE language features in their lyrics to construct an identity'.

Language investigation: hypotheses, research questions and data collection

- Notice the difference between the two examples above: whilst the first is phrased with some element of certainty (an educated guess based around linguistic methods), the second hints that you're not sure of the outcome, but will follow a linguistic framework to test your question.
- It is important to be precise and give your investigation a clear focus: whose language exactly are you investigating? Under what conditions or circumstances? What specific language features are you going to look at?
- Ensure that your hypothesis is based around some sort of language level. So if you're looking at the difference between male and female speech in the classroom, it's better to state that 'I think men will have a longer turn length or MLU' rather than 'I think that men will talk more'.
- You can make more than one hypothesis. But be wary of making too many, as you'll find it difficult to reference each hypothesis in enough detail in the analysis section of your investigation, and you'll lose marks.

Collecting data

Once you have decided on a topic and have an idea as to what you'd like to investigate, it's time to think about what sort of data you would need for your project. Of course, the type of data and the questions you are hoping to look into will determine the methods you use. There are many different things you can do, and you can use more than one type of data collection:

- **Spoken data** - you can record conversations or collect video clips (for example from YouTube) or podcasts and make a transcription.
- **Written and mixed mode data** - any kind of text you are interested in (children's written work, articles from newspapers, magazines, websites, text messages, emails, books etc.).

The above are all texts that you would analyse. However, you may wish to collect other types of data to help you find other interesting/different things to say about the actual language data:

- **Questionnaires** - ideal if you want to find out how a number of people think or feel about particular language issues or if you'd like to see how a group of random (or not so random) people pronounce words or use language.
- **Interviews** - if you want to get a better understanding of one or a few people's use of language or attitudes to it, it allows for more depth than a questionnaire; you can use an interview as a follow-up from a questionnaire, selecting the most useful or interesting people from the questionnaires for further interview.
- **Experiments** - you can ask participants to take part in a controlled situation in which you observe what happens.

Ethics

When you undertake research it is important that you do it in an ethical fashion. For example, **if you want to investigate child language, you must ask permission of the parents** to do so. If you want to record people speaking you need to explain why you are doing it and ask for permission. If you are writing about individuals, you need to keep their names out of your research, so that individuals cannot be identified.

Below are some tasks that help you put the info above into practice. Answer each question in as much detail as you can.

1. Below are some research questions proposed by A-Level English language students. For each one, say what is good and/or poor about them. As an extension you could consider what language levels would be appropriate to help answer the question.

- a. How do children learn to read?
- b. How do teachers use language to manage behaviour in classrooms?
- c. How is dyslexia affecting my little brother's writing?
- d. The language of song lyrics.
- e. What is the difference between the language of tabloid and broadsheet newspapers?

2. In the table below are some of the topics chosen by A-level English language students for their investigations. Have a go at formulating a research question and a hypothesis for each.

Suggestions:

- Think about the relevant language levels and if applicable relevant linguistic theories and concepts that would apply.
- Always keep your research questions as open as possible. Avoid setting a biased hypothesis such as 'What is the difference between masculine and feminine speech styles?' This might make you run the risk of looking for differences, ignoring similarities possibly and even explaining all findings in the light of genderlect, while other contextual issues may be more significant as explanations.
- As mentioned above, ensure that your hypothesis includes some reference to a language level. A language concept is also ok - for example you could include theorists such as Tannen or Labov in your hypothesis.

	Topic	Research question and/or hypothesis
1	The language used by a driving instructor in driving lessons.	
2	The language featured in packaging of healthier options for breakfast cereals.	
3	Comparing two four-year-olds from the same nursery class: one who is growing up as a monolingual English speaker and the other who is growing up as a bilingual English and French speaker.	
4	The ways in which national newspapers in the UK represent the most recent Prime Ministers: David Cameron and Theresa May.	
5	The language used by celebrities well known for their Twitter presence in their tweets as a means of creating and maintaining their celebrity personas.	

3. For each of the methods above, list advantages and any disadvantages. Discuss how the disadvantages could be minimised or overcome.

4. Here are some examples of research methods employed by students in the past. Can you explain what you think the student's focus for their investigation was? There may be more than one possible answer in some cases, so try to think of as many answers as possible:

- a. **Experiment:** asking speakers for directions in local area and then pretending to have misunderstood and ask again to observe whether speakers change any aspect of their pronunciation.
- b. **Questionnaire:** asking staff and students in college to give their opinions on commonly used examples of non-standard English e.g. 'we were sat'.
- c. Recordings of family members' interactions during dinner.
- d. Recordings of GP with patient in consultation room.
- e. A corpus of online newspaper articles featuring the term 'immigrant' following the 24 hours after the Brexit referendum in 2016.

Could you use or adapt these examples to suit your investigation topic or focus? How would you use the same research method to suit your chosen research question or hypothesis?

5. Which method(s) would you suggest to a student wanting to investigate the following (you can recommend more than method)?

- a. How a speaker of other language(s) is finding that their English is being affected by their other language(s).
- b. People's attitudes to a northern English accent.
- c. How teenage males and teenage females speak in single and mixed sex groups.
- d. Current taboo words (words considered offensive).
- e. The language of football commentators.

6. A student is really interested in exploring the representation of men and women in advertising. In order to obtain data, they decide to search online with the search term: 'sexist adverts' and select adverts suggested by the search engine.

What problems can you see with this method of data collection? List at least two problems and suggest better ways for this student to collect their data.

7. If you wanted to record speech, but you didn't want to speakers to be aware of the fact that they are being recorded (to avoid **observers' paradox** - where people are aware they're being recorded), how could you make sure to do this in an ethical way? Why might the observers' paradox be an issue when it comes to the accuracy of your data?

8. If you wanted to investigate an aspect of child language acquisition, how would you ensure that you obtain any spoken or written data ethically?