## READING PASSAGE 3

*You should spend about 20 minutes on **Questions 27–40** which are based on Reading Passage 3 below.*

# Obtaining Linguistic Data



**A** Many procedures are available for obtaining data about a language. They range from a carefully planned, intensive field investigation in a foreign country to a casual introspection about one's mother tongue carried out in an armchair at home.

**B** In all cases, someone has to act as a source of language data – an *informant*. Informants are (ideally) native speakers of a language, who provide utterances for analysis and other kinds of information about the language (e.g. translations, comments about correctness, or judgements on usage). Often, when studying their mother tongue, linguists act as their own informants, judging the ambiguity, acceptability, or other properties of utterances against their own intuitions. The convenience of this approach makes it widely used, and it is considered the norm in the generative approach to linguistics. But a linguist's personal judgements are often uncertain, or disagree with the judgements of other linguists, at which point recourse is needed to more objective methods of enquiry, using non-linguists as informants. The latter procedure is unavoidable when working on foreign languages, or child speech.

**C** Many factors must be considered when selecting informants – whether one is working with single speakers (a common situation when languages have not been described before), two people interacting, small groups or large-scale samples. Age, sex, social background and other aspects of identity are important, as these factors are known to influence the kind of language used. The topic of conversation and the characteristics of the social setting (e.g. the level of formality) are also highly relevant, as are the personal qualities of the informants (e.g. their fluency and consistency). For larger studies, scrupulous attention has been paid to the sampling theory employed, and in all cases, decisions have to be made about the best investigative techniques to use.

**D** Today, researchers often tape-record informants. This enables the linguist's claims about the language to be checked, and provides a



LIVEWORKSHEETS

way of making those claims more accurate ('difficult' pieces of speech can be listened to repeatedly). But obtaining naturalistic, good-quality data is never easy. People talk abnormally when they know they are being recorded, and sound quality can be poor. A variety of tape-recording procedures have thus been devised to minimise the 'observer's paradox' (how to observe the way people behave when they are not being observed). Some recordings are made without the speakers being aware of the fact – a procedure that obtains very natural data, though ethical objections must be anticipated. Alternatively, attempts can be made to make the speaker forget about the recording, such as keeping the tape recorder out of sight, or using radio microphones. A useful technique is to introduce a topic that quickly involves the speaker, and stimulates a natural language style (e.g. asking older informants about how times have changed in their locality).

**E** An audio tape recording does not solve all the linguist's problems, however. Speech is often unclear and ambiguous. Where possible, therefore, the recording has to be supplemented by the observer's written comments on the non-verbal behaviour of the participants, and about the context in general. A facial expression, for example, can dramatically alter the meaning of what is said. Video recordings avoid these problems to a large extent, but even they have limitations (the camera cannot be everywhere), and transcriptions always benefit from any additional commentary provided by an observer.

**F** Linguists also make great use of structured sessions, in which they systematically ask their informants for utterances that describe certain actions, objects or behaviour. With a bilingual informant, or through use of an interpreter, it is possible to use translation techniques ('How do you say *table* in your language?'). A large number of points can be covered in a short time, using interview worksheets and questionnaires. Often, the researcher wishes to obtain information about just a single variable, in which case a restricted set of questions may be used: a particular feature of pronunciation, for example, can be elicited by asking the informant to say a restricted set of words. There are also several direct methods of elicitation, such as asking informants to fill in the blanks in a substitution frame (e.g. *I ___ see a car*), or feeding them the wrong stimulus for correction ('Is it possible to say *I no can see*?').

**G** A representative sample of language, compiled for the purpose of linguistic analysis, is known as a *corpus*. A corpus enables the linguist to make unbiased statements about frequency of usage, and it provides accessible data for the use of different researchers. Its range and size are variable. Some corpora attempt to cover the language as a whole, taking extracts from many kinds of text; others are extremely selective, providing a collection of material that deals only with a particular linguistic feature. The size of the corpus depends on practical factors, such as the time available to collect, process and store the data: it can take up to several hours to provide an accurate transcription of a few minutes of speech. Sometimes a small sample of data will be enough to decide a linguistic hypothesis; by contrast, corpora in major research projects can total millions of words. An important principle is that all corpora, whatever their size, are inevitably limited in their coverage, and always need to be supplemented by data derived from the intuitions of native speakers of the language, through either introspection or experimentation.

www.luckyidiom.com



LIVEWORKSHEETS

## *Questions 27–31*

Reading Passage 3 has seven paragraphs labelled **A–G**.

Which paragraph contains the following information?

*Write the correct letter **A–G** in boxes 27–31 on your answer sheet.*

***NB** You may use any letter more than once.*

**27** the effect of recording on the way people talk

**28** the importance of taking notes on body language

**29** the fact that language is influenced by social situation

**30** how informants can be helped to be less self-conscious

**31** various methods that can be used to generate specific data

## *Questions 32–36*

www.luckyielts.com

*Complete the table below.*

*Choose **NO MORE THAN THREE WORDS** from the passage for each answer.*

*Write your answers in boxes 32–36 on your answer sheet.*

| METHODS OF OBTAINING LINGUISTIC DATA | ADVANTAGES | DISADVANTAGES |
|---|---|---|
| **32**..... as informant | convenient | method of enquiry not objective enough |
| non-linguist as informant | necessary with **33**..... and child speech | the number of factors to be considered |
| recording an informant | allows linguists' claims to be checked | **34**..... of sound |
| videoing an informant | allows speakers' **35**..... to be observed | **36**..... might miss certain things |



LIVEWORKSHEETS

## *Questions 37–40*

*Complete the summary of paragraph G below.*

*Choose **NO MORE THAN THREE WORDS** from the passage for each answer.*

*Write your answers in boxes 37–40 on your answer sheet.*

A linguist can use a corpus to comment objectively on **37**..... . Some corpora include a wide range of language while others are used to focus on a **38**..... . The length of time the process takes will affect the **39**..... of the corpus. No corpus can ever cover the whole language and so linguists often find themselves relying on the additional information that can be gained from the **40**..... of those who speak the language concerned.

www.luckyielts.com



LIVEWORKSHEETS