RESEARCH METHODS IN LINGUISTICS FOR THE STUDY OF POLITICAL DISCOURSE: AN OVERVIEW

Boryana Kostova[1]

Department of Foreign Language Teaching and Post-graduate Qualification,
Varna Free University

Abstract: The paper attempts to present an overview of research methods within linguistics. It starts by considering the main research paradigms and classification of research, discusses key research methodology issues and focuses on the current prevalence of certain methods in the field of linguistics. Without providing a comprehensive description and analysis of all the methods within the field, only a limited group of specific methods are discussed which have become highly influential in the study of political discourse. These research methods are introduced through their distinctive features, underlying principles and main contribution to linguistic research. Various empirical studies of political discourse are reviewed as illustration of how ethnography, discourse analysis, Critical Discourse Analysis, narrative analysis, questionnaire surveys, and corpus linguistics are applied as research methods. The main conclusions drawn relate to the potential and limitations of the methods as well as to the clear trends in their application when designing linguistic research of political discourse.

Key words: research methods, linguistics, political discourse

[1] **About the author:** Dr. Boryana Kostova is Senior Lecturer at the Department of Foreign Language Teaching and Post-graduate Qualification, Varna Free University, Varna, Bulgaria. She has extensive experience in teaching various ESP courses, materials design and certification of language proficiency. Her research interests and publications are in the area of political discourse, academic communication and ELT.

e-mail: boryana.kostova@vfu.bg **ORCID ID:** 0000-0002-6207-924X

Introduction

The study of research methods is an attempt to identify the activities that are characteristic of research. Awareness of research methods is important for several reasons. First, their insightful application is a manifestation of the researchers' competence. Also, accurate and reasonable methodology is not only an instrument but the road for investigation of a topic, verification of the results, generation of new ideas and theories, and persuasion of the communities of practice of the validity of the conclusions.

The starting point of the paper is the understanding that no fixed toolkit of methods common across all disciplines exists. Therefore an overview of research methods within linguistics is provided. Without attempting at a comprehensive description and analysis of all the methods within the field, only a limited group of specific methods are discussed. These are introduced by focusing on their key features, underlying principles and main contribution to linguistic research. In order to illustrate them the paper draws on linguistic studies of political discourse.

1. Theoretical Assumptions

1.1. Key concepts

The key concepts central to this study are *method* and *research*. *Method* can be defined by distinguishing it from (Hepburn and Andersen, 2021):

1. the aims of research, such as knowledge, prediction, control, etc.
2. methodology, which includes the values associated with a particular research method, such as objectivity, reproducibility, simplicity, etc. Methodology is rather a normative theory of how research should be conducted to generate knowledge without indicating the particular methods (Wodak and Meyer, 2016, p. 16).
3. practices through which methods are implemented, such as specific laboratory techniques, ways of communicating the results, habits and accepted ways of conducting research.

So, the term *method* refers to a set of techniques and practices for practical and theoretical investigation of reality. It is associated with "techniques for gathering evidence, e.g. for collecting and selecting data, but also for explaining relationships, for conducting interpretations in a transparent way" (ibid.).

The term *research* is traditionally defined as a systematic approach to answering questions. In the scientific sense Dörnyei (2007, p. 15) refers to it as a *disciplined inquiry* giving a useful shorthand for the organized, systematic search for answers to the questions we ask. Another more detailed definition should also be taken into consideration when discussing what research is:

> [R]esearch in applied linguistics can be defined as an investigation, examination or inquiry that requires planning, organizing and ethical considerations as well as systematic and careful analysis of data, sound interpretations and conclusions on the basis of evidence and inferences being made.
>
> (Phakiti, Paltridge, 2015, p. 22)

However, these definitions of *research* tend to be controversial due to the different interpretations of concepts such as *question, systematic approach* and *answer* (Farhady, 2013, p. 5196). The differences stem from the different theoretical orientations scholars adhere to. Obviously research methods are governed by some abstract principles. Therefore, the first step in their discussion is to provide an explanation of their main philosophical underpinnings.

1.2. Research paradigms

Every research is performed within a particular school of thought. This underlying philosophical view of what is considered knowledge or reality is defined as *research paradigm*. The three dimensions of a research paradigm are: ontology, epistemology and methodology (Phakiti, Paltridge, 2015, p. 27). *Ontology* refers to the nature of reality and the nature of human beings in the world; *epistemology* relates to the nature of knowledge and how it can be achieved; *methodology* means the research approach that is used to investigate reality. Farhady (2013) suggests that *axiology* or the ethical stance towards the world and the self of the researcher is another aspect of research paradigms. Although there are numerous research paradigms, it is widely agreed that the main research paradigms within linguistics are positivism, post-positivism, constructivism and pragmatism (Dörnyei, 2007, pp. 9-11; McKinley, 2020, p. 3; Phakiti, Paltridge, 2015, pp. 27-29).

Positivism is a research approach according to which reality can be understood objectively as the object of research exists objectively. It is believed that there are laws or theories that rule and explain physical or social reality. These laws do not depend on the personality of the researcher, the time or the way reality is examined. Reality, on the other hand, is seen as quantifiable and measurable. Therefore, positivism is the basis of quantitative research. In the search of objectivity and an accurate correspondence between their observations and the reality, the focus is put not on the research setting. Rather the research is conducted by defining and controlling variables and manipulating the research setting. Research must objectively test a hypothesis by using scientific method and/or logic to prove it to be true.

Post-positivism is a modified version of positivism. Unlike positivism, it argues that complete objectivity is almost impossible to achieve and acknowledges the researcher's subjectivity. It believes that although the object of inquiry exists independently of the human mind, it cannot be observed or perceived with total accuracy. Still, objectivity is considered as an ideology which should guide researchers and their research.

Unlike positivism and post-positivism, *constructivism* or *interpretivism* is a research approach that views social realities as multiple and dependent on the researcher and participants in the study, the nature of the study and the context in which the study is conducted. This is a relativist and a subjectivist position. Because the subject matter of social sciences differs from the subject matter of physical or natural sciences, it requires a different aim for study and different methods. On the other hand social reality, i.e. cultures, institutions and values, cannot be reduced in the same way as physical reality. These are seen as socially co-constructed. The perceived reality can be multiple because reality is seen as dependent upon the researcher's values and background. Thus research is built on the idea that knowledge is actively constructed, usually through human interpretation of experience. This is the reason why non-experimental, non-manipulative research procedures are used.

Pragmatism is a pluralistic approach to research. It does not view truth or reality as something independent of human mind or socially co-constructed. Objectivity and subjectivity are considered only in relation to the research topics. Pragmatism focuses on the importance of problem solving, i.e. a research method should be chosen according to its effectiveness in answering the research question. Thus the application of a variety of methods from various research paradigms is allowed, as well as different types of data and data analyses. As it attempts to use the best of both quantitative and qualitative research methods to maximize understanding of the research problem, pragmatism is the research paradigm that underpins mixed methods research.

2. Types of research

2.1. Main classifications of research

There are various types of research. The main classifications focus on the distinction between primary and secondary research, basic and applied research, cross-sectional and longitudinal research and quantitative and qualitative research. Mixed methods are another significant research approach (Dörnyei, 2007, pp. 19-21; Phakiti, Paltridge, 2015, pp. 23-26).

The distinction between *primary* and *secondary* research relates to the data needed to answer a research question. In *primary research* empirical data is collected to answer the research questions. The research uses particular instruments or techniques to collect data which is then analysed in order to answer the research questions.

Secondary research relies on other research and sources to answer the research questions. It can also be described as *library study* or a *review of the literature* on a particular topic. This form of research can include existing theories or current knowledge on a particular issue that can then be used to form the basis of a primary research project. Another form of secondary research has become known as *research synthesis*, which includes a *meta-analysis* that aims to gain information about the current status of knowledge through a systematic analysis of empirical findings in a particular research area. This form of systematic review is useful as it is able to inform researchers as to what still needs to be done in future research.

The main distinction between *basic* and *applied* research lies in what researchers aim to achieve in their research. *Basic research* aims to produce fundamental knowledge so that explanations of a phenomenon are meaningful and sufficient. *Applied research* is related to situations in which researchers or practitioners aim to use or apply knowledge or theories from basic research to address a problem. This is conducted by systematically applying the theories. For example activities are done with a group of individuals and observation is made to conclude how they work in order to improve the learning process. In applied research, researchers aim to seek solutions to a problem. Most studies in applied linguistics contain elements of both basic and applied research.

The distinction between *cross-sectional* and *longitudinal research* relates to the *time* at which the data is collected and the *length* of time taken for the data collection. *Cross-sectional research* refers to research in which researchers collect data from one or more cohorts (a person, group of people) at a single point in time or within a short period of time by using questionnaires. Cross-sectional research is often described as a *snapshot of data collection*. Survey research often adopts a cross-sectional data collection design. *Longitudinal research* refers to research in which researchers gather the same aspects of information from the same participant(s) over a period of time. This allows researchers to observe changes or stability in behaviours, learning, abilities, cognitive and social development. Longitudinal research can help researchers establish patterns or changes.

The dichotomy between *quantitative* and *qualitative research* is one of the most significant distinctions in research methodology. As pointed out by Phakiti and Paltridge (2015, p. 24) it primarily refers to the types of data collected by the researchers – using figures versus non-quantitative data. Whereas quantitative research involves data collection procedures that result mainly in numerical data that is later analysed by statistical methods, qualitative research involves data collection procedures that result primarily in open-ended, non-numerical data which is analysed by non-statistical methods. Obviously, the terms *quantitative* and *qualitative research* reveal the general ideological orientation of the study, the method of data collection, the nature of the collected data, and the method of data analysis (Dörnyei, 2007, pp. 24-29). It can be concluded that the contrast between quantitative and qualitative research is based on three sources: ideology, categorization and perception of individual diversity. First, the terms qualitative and quantitative point to an ideological contrast because qualitative inquiry was created as a reaction or a reformist movement against the mainstream quantitative paradigm. Second, the two terms name differing categorizing practices. While quantitative research uses a predetermined numerical category system, qualitative research uses flexible verbal coding. Third, quantitative research uses large samples to exclude individual variability both at the data collection and data analysis stages. Qualitative research, on the other hand, is characterized by sensitivity to the individual as it is linked to two sources of variation, i.e.zx the individual respondents and the individual researchers with their subjective sensitivity, training and experience.

Although the two types of research are sometimes metaphorically represented as a *paradigm war*, they are not exclusive, but rather form a continuum and lay the foundations of a third research approach. The integration of the two research methodologies underpins mixed methods research. This involves different combinations of qualitative and quantitative research, which can be used either at the data collection stage or at the analysis level.

2.2. Quantitative research

Quantitative research is one of the most widely practiced approaches in social sciences in general, and in applied linguistics in particular (Farhady, 2013, p. 5196). In the social sciences it has been used since the nineteenth century when researchers adopted the so called *scientific method* in their investigations inspired by the progress in natural sciences (Dörnyei, 2007, p. 30-34). The *scientific method* was closely related to numerical values and statistics. It consists of three main stages: (1) observing a phenomenon or identifying a problem; (2) generating an initial hypothesis; (3) testing the hypothesis by collecting and analyzing empirical data using standardized procedures. When the hypothesis is tested and validated, it is accepted as a scientific theory or law. This is a way to conduct a study objectively, to minimize the subjective influence of the researcher and consequently to have an accurate and reliable description of the world. While quantitative methodology became dominant in the social sciences in the first half of the twentieth century, it was only in the 1970s that it started to characterize research in applied linguistics.

Quantitative research in the social sciences has the following characteristic features:

1. It is centred around numbers. In order for the numbers to work, they need contextual backing, i.e. precise definition of the content and the boundaries of the variables as well as exact descriptors for the range of values that are allowed within a variable.
2. Categories and values are specified prior to the actual study. A quantitative study requires a long and careful preparation and piloting.
3. Quantitative research focuses on the study of variables. The aim is to identify common features through counting, scaling or assigning value to data. Another aim is to identify the relationships between variables by measuring them and manipulating them.
4. Using statistics and statistical terminology is the most salient feature of quantitative methods.
5. Quantitative methodology uses standardized research procedures which remain the same throughout the study. The rationale behind this is to eliminate human variability and bias. This emphasis on objectivity guarantees that the findings describe the objective reality independent of the researcher's subjectivity.
6. The aim of quantitative research is to generalize and to find out universal laws.

As pointed out by Dörnyei (2007, pp. 34-35) quantitative findings are almost universally accepted. The reason is that quantitative studies are systematic, rigorous, focused, use precise measurement, their analytical apparatus is refined, possess in-built quality checks and indices. All this leads to the validity of the findings. Another strength is their quick research process which relies on statistical software. Nevertheless there are some disadvantages like averaging responses across the observed group, not paying attention to the subjectivities, as well as lacking sensitivity towards the reasons for particular observations.

2.3. Qualitative research

Qualitative research developed as a reaction to the shortcomings of quantitative research which was considered simplistic, decontextualized, failing to relate to real human life and circumstances when analysing complex phenomena. Dörnyei (2007, pp. 35-42) traces the beginning of qualitative research as early as 1900s within social sciences. In applied linguistics there has been a growing interest in the approach since 1990s. According to Benson (2013, p. 5195) qualitative methods in applied linguistics appear both in the form of established approaches borrowed from the social sciences and in the form of more eclectic use of methods and techniques. Their most salient features are the following:

1. Qualitative studies are open to new details that may emerge during the process of investigation. This flexibility refers both to the research questions, which may change or be refined during the study, and the researchers, who enter the research process without preliminary hypotheses. This means that the analytic categories are defined during the process of the research.
2. Qualitative research works with a wide range of data – recorded texts, texts from various genres, images. A. Holliday (2015, p. 49) has observed that "the aim of qualitative research is to search for the richest possible data". Qualitative data is not gathered with the aim to be directly counted or measured in an objective way. The purpose of data is to reveal complex details and cultural and personal meanings.
3. Qualitative research is conducted in a natural setting, without any attempts to manipulate it.
4. Qualitative researchers aim at viewing social phenomena from the perspective of the insiders. This relates to the belief that human behavior is based upon meanings that people

attribute to and bring to situations. Qualitative research is connected to the constructivist view that each individual creates his or her own unique understanding of the world, so there are multiple constructions and multiple interpretations of reality (Benson 2013, p. 5190).

5. Qualitative studies use much smaller samples of participants than quantitative ones.
6. Qualitative research is the product of the researcher's subjective interpretation of the data. Holliday (2015, p. 50) also acknowledges that "the ideas and presence of the researcher is influential in what the data looks like and the way it is interpreted. […] The outcomes of the research will always be influenced by the researcher's beliefs".

However, qualitative research has a few weaknesses compared to quantitative research.

- It uses very small participant samples and approach the issue of generalizability in a different way. This means that sometimes there is certain over-interpretation of the individual cases and the conclusions may not be universally applicable.
- It may be influenced by the researcher's personal biases and individual characteristics.
- Because qualitative research does not use standardized instruments and procedures and statistical techniques, it may appear unprincipled and anti-methodological.
- It may build too narrow theories out of individual cases studied.
- Qualitative research and data processing involves investment of much more time and effort than quantitative research.

2.4. Mixed methods research

The integration of qualitative and quantitative methods in a single research project is referred to as mixed methods research. It is recognised as the third approach in research methodology. It involves the use of multiple methods and the collection of multiple forms of data. Although the practice of mixing very different research approaches has had a long history outside social sciences, mixed methods research has been in use in applied linguistics since 1990s and has been directly associated with two paradigms – transformative and pragmatic (Ivankova, Greer, 2015, p. 63). The transformative paradigm is guided by the principles of social justice, whereas pragmatism argues that what has practical and functional value is ultimately important and valid.

Explicit discussion of mixing methodologies started only in 1970s with the introduction of the concept of *triangulation* (Dörnyei 2007, pp. 43-45). The term *triangulation* entered the social sciences when it was borrowed from naval navigation and land surveying. Originally it referred to a method for determining the yet unknown position of a certain special point through measurement operations from two unknown points. In social research it became synonymous to combining data sources to study the same social phenomenon. There is more than one type of triangulation (Angouri, 2010, p. 34):

- Data triangulation (the application of more than one sampling method for data collection);
- Investigator triangulation (the involvement of more than one researcher);
- Theoretical triangulation (the use of more than one theoretical stance);
- Methodological triangulation (the use of more than one methodology).

Methodological triangulation is seen as a way to reduce the inherent weaknesses of individual methods by emphasizing the strength of the other and an instrument for increasing the internal and external validity of research. Triangulation is considered as the main purpose of mixed methods research (Phakiti, Paltridge, 2015, p. 25).

Another important issue that must be taken into consideration is the difference between *data triangulation* and *mixed methods research*. Although a mixed methods design may allow researchers to triangulate research results, data triangulation refers only to the strategy of collecting information from different or multiple sources to help gain a deeper understanding of a particular matter. For example, ethnographic researchers who use a combination of interviews, observations and document analysis to answer their research questions are not necessarily employing a mixed methods design since they analyse the data qualitatively, not mixing quantitative and qualitative methods. Similarly, experimental researchers who collect data from language tests, academic grades, self-assessment and various kinds of questionnaire are not necessarily using a mixed methods design since they may choose to analyse all the data statistically. In both examples, data triangulation is not the same as mixed methods research. (ibid., p. 26)

Apart from triangulation, or corroboration of results from different methods, there are other common mixed methods designs in applied linguistics research which depend on the purposes of the research (Angouri, 2010, p. 35; Ivankova, Greer, 2015, p. 61; Phakiti, Paltridge, 2015, p. 27):

(1) Researchers may aim at discovering meaningful contradictions and the ‘paradox’. This research design is called *initiation*. It refers to the discovery of new perspectives from one method with questions or results from the other method.

(2) Researchers can use one research method to complement another. This is a one-after-the-other design or *complementarity*. The study may begin with a research question that can be answered through quantitative analysis. On the basis of the findings of the quantitative component of the study, qualitative data may be collected. The researchers then analyse the qualitative data to add to or counterbalance the quantitative findings.
(3) Researchers can use one research approach as a starting point for another. The reason for the use of the approach is *development*, i.e. to use the results from one method to help or inform the other method. This is a dominance design where the emphasis is more often on the latter method of the two.
(4) Researchers can ask a set of research questions and use one method to answer each question separately and independently. This is referred to as a side-by-side design or *expansion*. The reason for the use of mixed methods is to broaden the scope and objective of the research as well as the range of inquiry by using different methods for different inquiry components.

Based on the variation of the methodological characteristics discussed above, three basic types of mixed methods research designs can be used in applied linguistics (Ivankova, Greer, 2015, pp. 66-69):

(1) *Concurrent Quantitative + Qualitative mixed methods research design* is used when it is necessary to compare or merge quantitative and qualitative results to produce well-validated conclusions. It typically includes two strands, during which quantitative and qualitative data are collected and analysed separately or independently of each other. The mixing of the quantitative and qualitative methods occurs when quantitative and qualitative results are compared or synthesized to find evidence and to have a more complete understanding of the research problem.
(2) *Sequential Quantitative → Qualitative mixed methods research design* is used when there is a need for follow-up qualitative data to explain or confirm initial quantitative results. It consists of two chronological strands with a quantitative strand occurring first in sequence. The weight is typically given to the first, quantitative, study strand because this design is mostly used when the research problem and related purpose require examination by quantitative methods. The mixing of the quantitative and qualitative methods typically occurs chronologically at the completion of the first, quantitative, strand and beginning of the second, qualitative, strand.

(3) *Sequential Qualitative → Quantitative mixed methods research design* is used when it is necessary to use initial qualitative data to develop new measures and identify unknown variables and relationships. It consists of two chronological strands with a qualitative strand occurring first in sequence. The weight is typically given to the first, qualitative, study strand, because in this design the researcher starts by qualitatively exploring a topic before building a second, quantitative phase. The mixing of the qualitative and quantitative methods typically occurs chronologically at the completion of the first, qualitative, strand and the beginning of the second, quantitative, strand.

Despite of all these issues, the popularity of mixed methods research is growing. Here are the arguments for this (Angouri, 2010, pp. 29-33; Dörnei, 2007, pp. 45, 166; Parmelee, Perkins, 2013, p. 1970):

1. A mixed methods approach is valuable because every method has its limitations. Each method is designed to find answers to certain types of research questions. As a result, using only one method can limit the ability of the study to answer all the research questions comprehensively. Thus the combination of qualitative and quantitative methods can improve the quality as the strengths of one method can be used to overcome the weaknesses of another method used in the study.
2. A better understanding of a complex phenomenon can be achieved from triangulated findings. Combining the two paradigms is beneficial for giving answers to a various research questions and can point to different layers of meaning.
3. The validity of the research outcomes is improved through the convergence of findings.
4. The final results are acceptable for a larger audience than those of one-dimensional study because the findings are rich.

Although this is a potentially progressive approach, there are certain weaknesses of using mixed methods research. The most obvious one is the fact that researchers are not usually adequately trained in both methods (Dörnyei, 2007, p. 46). Another danger arises from the belief that mixing methodologies represent a collection of techniques, an unstructured mix of quantitative and qualitative research or just the sum of both paradigms. This does not necessarily lead to better research. The main concern is the amount of integration of the two paradigms (Angouri, 2010, p. 31-33).

3. Research methodology issues

3.1. Quality criteria for research

Methodological decisions depend not only on the choice of research paradigm. Another issue to be considered is the quality criteria for research because explicit quality standards need to be achieved (Dörnyei, 2007, pp. 48-62). One problem is that no universally accepted terminology to describe quality criteria exists. However, the terms which are relevant and widely used in both qualitative and quantitative research are *validity* and *reliability*. Both terms were originally introduced in quantitative research. Within qualitative research other alternative terms like *trustworthiness*, *authenticity*, *credibility* and *rigour* have also been used.

Quality criteria in quantitative research can be divided into three parts: reliability, measurement validity and research validity. In the field of Applied Linguistics *reliability* indicates the extent to which the measurement instruments and procedures produce consistent results in a given population in different circumstances. The variation of the circumstances may involve differences in administrative procedures, changes in participants, differences in various forms of the test and differences in raters. If these variations cause measurement errors, then the results are unreliable. Reliability is not a characteristic of the instrument but a property of the scores on a test for a particular population of test-takers. *Research validity* concerns the whole research process and focuses on the distinction of (a) internal validity, or meaningfulness of the research, and (b) external validity, or the generizability of the results beyond the observed sample, meaning that results can apply beyond the unique sample. *Measurement validity* refers to the meaningfulness and appropriateness of the interpretation of the various test scores or other assessment procedure results.

Setting explicit quality criteria in qualitative research has been difficult. One problem is that the terms *validity* and *reliability*, associated with empirical research, have been appropriate for quantitative methods. The problem arises from the fact that qualitative study is by nature subjective, interpretative and dependent on context. In this this understanding truth is relative while facts depend on individual perceptions. Thus within Applied Linguistics qualitative research *reliability* refers to the degree of consistency with which instances are assigned to the same category by different observers or by the same observer on different occasions. Instead of validity the concept of *trustworthiness* is introduced. According to Lincoln and Guba's taxonomy of quality criteria *trustworthiness* consists of four components (ibid., p. 57):

1. Credibility which corresponds to internal validity in quantitative research;
2. Transferability, or applicability of the results to other contexts, which corresponds to external validity in quantitative research;
3. Dependability, or consistency of the findings, which corresponds to reliability in quantitative research;
4. Confirmability, or neutrality of findings, which corresponds to objectivity in quantitative research.

The standards of quality for qualitative research can be viewed in terms of Maxwell's taxonomy of validity in qualitative research. It consists of five components (ibid., p. 58):

1. Descriptive validity, or the factual accuracy of the researcher's account. A strategy for ensuring this primary validity type is to use multiple investigators to collect and interpret data.
2. Interpretative validity. It focuses on the quality of the portrayal of participants' perspective. A strategy to achieve this type of validity is to discuss the findings with the participants.
3. Theoretical validity. It refers to the appropriate level of theoretical abstraction and how well the selected theory explains or describes the phenomenon in question.
4. Generizability. It has two aspects – internal and external generizability. These aspects respectively refer to generalizing within or outside the community observed. A strategy to meet this criterion is to include participants' own judgements about the generizability of the studied issue.
5. Evaluative validity. It refers to the way the researcher evaluates the issue studied in terms of usefulness, practicability, etc.

In relation to quality criteria in mixed methods research it must be clear that the specific methods that are combined are either qualitative or quantitative. Thus the quality principles of these methods apply to mixed methods research.

3.2. Research questions and hypotheses

Research methods, as emphasized by Litosseliti, are “inextricably linked with the research questions being asked, as well as with the broader research climate in which they are employed” (2010, p. 2). Dörnyei (2007) also acknowledges that the proper way to do research involves two main stages: first, generating research questions; and then choosing the design, the method and the instruments for answering these questions. This means that the data needed, data collection methods and data analysis are all guided by the research questions of any empirical research project.

Research questions can be divided into several types (Sunderland, 2010, pp. 12-14):

- descriptive/explanatory/evaluative;
- primary/secondary;
- main/contributory;
- overarching/subordinate;
- empirical/methodological/theoretical;
- researcher-generated/participant-generated;
- empirical/speculative.

The type and number of research questions included in a study depends on the scale of the project. Whatever their volume and variety is, they should constitute a coherent whole, i.e. be explicitly related to each other. They also have to be clearly formulated, intellectually worthwhile and researchable because these properties relate to the originality of the research and its contribution to knowledge. All research questions should be referred to in the discussion. The expectation is that they are all addressed, discussed and their implications identified.

There is a difference between a research question and a hypothesis (ibid., p. 10). While these two notions are related, a hypothesis is more characteristic of natural than social sciences and tends to be more precise. A hypothesis is conventionally worded as a statement, which is to be investigated and proved or disproved through empirical study. Hypotheses are more characteristic of quantitative research while research questions are characteristic of qualitative research and are likely to be both broader and more exploratory than hypotheses.

Research questions and hypotheses have different roles in the different research paradigms (Dörnyei, 2007, pp. 72-74). The starting point of every research is the research topic which broadly corresponds to research interest. In order to be addressed, it has to be developed into research purpose. This is a short statement of the objectives of the planned study as well as the reasons and significance of the research. The next step is formulating specific research questions. Their role is to convert the research purpose into specific questions that the planned study will aim to answer. Research questions are an important part of the research because they indicate readiness for the actual research understood as orientation to the research methodology that can best achieve the research purpose.

Qualitative and quantitative studies differ in terms of how the purpose of research is specified and how it is broken down into specific research questions. In quantitative studies the purpose statements identify the target variables and the causal and descriptive relationships between them. The research questions specify concrete methodological procedures. Research hypotheses are also drawn and contain the researcher's predictions.

In qualitative studies, due to their emergent nature, research purposes and questions are rather vague. The research purpose often contains only the specification of a situated phenomenon or a central idea that will be explored in order to develop new insights or a theory. The research questions tend to be broader. What is emphasized instead is the exploratory nature of the study.

In mixed methods studies specifying the research purpose and questions may be challenging. As suggested by Dörnyei (ibid.) an effective strategy is to start with a purpose statement, followed by a rationale for the specific mixed design applied and the choice of particular methods, and conclude by separate research objectives. The research questions can be presented at the introduction of each phase of the study.

3.3. Sampling and representation

Sampling refers to the selection of data in terms of the scope and depth of coverage. The rationale of sampling differs within quantitative and qualitative methods.

Sampling in quantitative research involves taking a smaller amount of data which is considered to be representative in some way of the larger population. As explained by Dörnyei "the sample is a subset of the population that is representative of the whole population" (2007, p. 96). Representativity refers to the ability of a random sample to mirror the structure of the totality of data (Reisigl & Wodak, 2016, p. 61). The accurate way a particular sample represents the larger population is a key criteria for downsizing because it ensures that the results are generalizable.

Sampling strategies can be divided into two groups: (i) probability sampling, and (ii) non-probability sampling. *Probability sampling* involves the following complex and expensive procedures: (a) *random sampling* where participants or data are picked at random; (b) *stratified random sampling* by which data are selected on the basis of preset criteria in order to reflect the make-up of the population more closely; (c) *systematic sampling* where the sample is obtained by selecting elements at regular intervals; (d) cluster sampling where larger units of the population are randomly selected. Probability sampling is a tough process and is an unrealistic aim within applied linguistics (ibid., p. 98). Therefore most actual research in applied linguistics uses non-probability samples although it is a compromise to use them. *Non-probability sampling* strategies include: (a) *quota sampling* and *dimensional sampling*; (b) *snowball sampling*; (c) *convenience* or *opportunity sampling*. The optimal sample size depends on the type of quantitative research.

Sampling or downsizing of the bulky collected data in qualitative research and in discourse analysis particularly is carried out according to specific criteria. These include frequency, (proto)typicality, intertextual and interdiscursive scope, salience, uniqueness and redundancy (Reisigl & Wodak, 2016, p. 39). The issue of representativity as a statistical criterion is not fully relevant for many discourse studies. Rather, the term 'representative data' is used in the sense of 'typical cases' within a more or less well-defined corpus. According to Reisigl & Wodak the concept of 'representativity' "is not operationalized in a clear way, and thus of minor importance for the majority of empirical research on discourse" (ibid., p. 61).

In applied linguistics, qualitative inquiry is also not concerned with how representative the respondent sample is or how the experience is distributed in the population. Thus the main goal of sampling is to provide diverse responses and deep understanding of the issue explored (Dörnyei, 2007, pp. 125-129). In order to achieve this the participant selection process should remain open for a long time until saturation is reached, or the point when additional data simply repeats the revelations of the previous informants. In order to achieve relatively quick saturation three specific sampling strategies can be applied: (1) *homogeneous sampling*, i.e. selecting participants with similar characteristics; (2) *typical sampling*, i.e. participants sharing typical features in relation to the research focus; (3) *criterion sampling*. Sampling strategies may focus on atypical, unique or extreme cases. These include: (i) *maximum variation sample*; (ii) *extreme or deviant case sampling*; (iii) *critical case sampling*. Another set of sampling strategies lacks theoretical underpinning but is very popular: (a) *snowball* or *chain sampling*; (b) *opportunistic sampling*; (c) *convenience sampling*. Whatever sampling strategies are chosen they have to be purposeful and in accordance with the object of study.

4. Specific Methodologies in Linguistic Research

Research methods, both quantitative and qualitative, can further be classified in terms of data collection and data analysis. Quantitative data collection methods comprise questionnaire surveys, experimental and quasi-experimental studies and collecting quantitative data via the Internet. Qualitative data collection methods include ethnography, interviews, focus group interviews, introspective methods, case studies, diary studies and research journals (Dörnyei, 2007, pp. 95-160). Quantitative data analysis methods encompass descriptive statistics, t-tests, analysis of variance, correlation, non-parametric tests, computerized data analysis while qualitative data analysis methods embrace qualitative content analysis, Grounded Theory, computer-aided qualitative data analysis, Discourse-Analytical Approaches (ibid., pp. 197-262). The data collection and data analysis methods can be seen as more general approaches and as more specific data creation methods. However, within the field of linguistics such a dividing line is difficult to draw. In practice there is considerable overlap between the data collection methods and the data analysis procedures, depending on the study (Litosseliti, 2010, p. 3).

4.1. Applying Qualitative Methods

The paper considers specific qualitative methods first. The importance presupposed by their precedence over quantitative methods in the paper as well as the larger amount of information under qualitative methods reflects the considerable prevalence of such methods currently in the field of linguistics. This is acknowledged by Litosseliti (ibid., p. 2).

4.1.1. Ethnography

Ethnography or linguistic ethnography, a term used when the method is used for data collection and mixed with other methods for linguistic analysis, is a theoretical and analytical framework which in terms of its epistemology stems from social constructivist and post-structuralist approaches. It also draws widely on work in linguistic anthropology (Creese, 2010, p. 138). Its origins lie in research by sociocultural anthropologists in the first decades of the twentieth century, and particularly in the "participatory observation" method (Levon, 2013, p. 196). The method postulates that the goal of social science research is to understand how the behavior observed within a community is linked to the beliefs and interpretive practices characteristic of that group. Thus ethnography is about generating a theory of social behavior that is based on both an insider and a more outsider perspective. Because phenomena are studied in their natural setting in terms of the meanings people bring to them, "ethnography's home is within qualitative research" as observed by Starfield (2015, p. 112).

The most obvious feature of ethnographic study is that it involves 'fieldwork' or participant observation in a natural setting. Other features include: focusing on participant meaning, i.e. events are looked at through the subjective interpretation of the participants; long experience of the researcher in the envioronment under investigstion; emergent nature, i.e. the exact focus of the research develops out of or due to the context (Dörnyei, 2007, p. 131). Ethnography uses various, even eclectic data collection techniques, including participant and non-participant observation, ethnographic interviews and structured questionnaires, ethnographer's own diary with field notes, journal entries and self-recordings, film and audio recordings, visual imagery, authentic documents, community artifacts (Dörnyei, 2007, p. 130; Levon, 2013, p. 204).

Current ethnographic research, as pointed out by Dörnyei (ibid.), aims at describing and analyzing the practices and beliefs of cultures, which are not necessarily understood as ethnic groups but as organizations, programmes and distinct communities of practice. Ethnography describes in detail the daily life of a community and the cultural meanings and beliefs the participants attach to their activities. Krzyżanowski (2018, p. 180-181) has observed that in the recent years ethnography has broadened the research philosophy and the scope of its techniques and methods. From a particular method of data collection preoccupied with 'distant' cultures, it has become a style of research that is distinguished by its objectives, which are to understand the social meanings and activities of people in a given field or setting. Ethnography encompasses political, organizational, policy-making analysis and studies such contexts as medical and courtroom settings and educational institutions. The new perspective, referred to as 'reflexive ethnography', links context-sensitive explorations with the context-specific dynamics. This means that ethnography recognizes the fluidity, complexity and inherent diversity of the explored social fields, which are not treated as settings anymore but as contexts. It also accepts the fact that processes and phenomena studied and observed in the course of ethnographic fieldwork may have evolved both within and beyond the studied groups and may have been motivated by the social power structures.

Ethnography, and its contemporary strand linguistic ethnography in particular, is characterized by disciplinary openness, eclecticism and interpretative stance. By its nature, it is an interdisciplinary research. For example, *Systemic Functional Discourse Analysis* could be used alongside ethnography when describing speech acts by focusing on discourse features such as 'follow-up moves' and 'question-answer sequences' and to develop an understanding of how language functios as a context. *Conversation Analysis* could similarly be used with linguistic ethnography to reveal the subtle shifts and sequences in the social organization of 'conversation', or 'talk-in-interaction' (Creese, 2010). Another example, given by Krzyżanowski (2018), is the Discourse-Ethnographic Approach which integrates the recent developments within ethnographic and critical-analytic research. It combines ethnography and Discourse-Historical Approach particularly as complementary general frameworks. Due to triangulating between a set of stages of analytical research, the different aspects of the ethnographic and discursive analyses can be balanced. Thus the research design in Discourse-Ethnographic Approach comprises three stages: (a) problem definition and relevant research questions, theorisation and pre-contextualisation; (b) fieldwork including contextualisation; (c) Discourse-Historical analysis.

According to Creese (2010, pp. 138-139) and Krzyżanowski (2018, p. 182-183) the combination of linguistics with ethnography – and their different analytical tools – offers a greater set of resources than each field of study could offer on its own. The main argument for undertaking this method is that ethnography can benefit from the analytical frameworks provided by linguistics, while linguistics can benefit from the processes of reflexive sensitivity required in ethnography. Ethnography is expected to be improved by the detailed technical analysis which linguistic brings, while linguistics is supposed to be intensified by attention to context. Ethnography provides linguistics with a close reading of context not necessarily represented in some kinds of interactional analysis, while linguistics provides an authoritative analysis of language use not typically available through participant observation and the taking of field notes.

The main concerns associated with the method are expressed by Dörnyei (2007, p. 133) and Creese (2010, p. 150):

- Ethnographic studies require extensive time investment.

- Language and culture have already been studied by linguistic anthropology. The term *linguistic ethnography* might describe studies of the same phenomena under a new name. The move forward may be associated with developments in the conceptualisations of context, culture, community and language in order to be in line with the terminological changes in the field.

Three empirical studies of political discourse are discussed as illustration of how ethnography is applied as a research method:

(1) Duranti's participant observation of a US senator's election campaign trail raised awareness about the many discursive practices and persuasive devices required to keep on track such a huge campaign and related persons (2006). The study, based on data collected during a year-long study of a Congressional campaign, identified three discursive strategies candidates use to maintain a coherent presentation of self while facing the demands of the political process.

(2) Krzyżanowski's account of the differences in the organizational behavior of representatives of different types of institutional bodies of the European Union starts with observation of their meetings (2018). The analysis of discourses about organizational practices in the EU is based on four types of collected empirical data, i.e. interviews, observations, documents and external sources, and uses the Discourse-Historical Analysis techniques. The combination of the two research methods within Discourse Ethnographic Approach proves to be relevant in studies which are problem-oriented and examine complex social and political contexts and power relations.

(3) Wodak's exploration of politics on 'backstage' or the everyday life of politicians is conducted in order to study the routines of international organizations, the European Parliament in particular (2012). Ethnographic research was needed, such as participant observation in organizations, in-depth and narrative interviews, shadowing of insiders to grasp the processes of political strategizing and decision-making. The analysis of the discursive practices of the organization was based on a case study and drew on the instruments of Critical Discourse Analysis and Discourse-Historical Analysis.

4.1.2. Discourse-Analytic Approaches

Current research in linguistics which relies on qualitative methods considers discourse analysis of particular importance. Discourse analysis is conducted through diverse analytic approaches like Conversation Analysis, Critical Discourse Analysis, Feminist Post-structuralist Discourse Analysis as well as Speech Act Theory, Interactional Sociolinguistics, Ethnography of Communication, Pragmatics, Variation Analysis, Discursive Psychology (Baxter, 2010). There is another set of methods which are also valuable for current linguistic research such as Content Analysis, Rhetorical and Argumentation Analysis, Narrative Analysis, Genre Analysis, Multimodal Analysis, Cognitive Linguistic Critical Discourse Studies (Wodak & Forchtner, 2018). Space constraints do not permit the provision of a comprehensive overview of all the discourse-analytical approaches. Therefore the paper focuses only on those which have become highly influential in the study of political discourse.

4.1.2.1. Discourse Analysis

Discourse Analysis or Discourse Studies refers to a wide range of methods that are useful in answering research questions in the fields of linguistics, humanities and social sciences (Johnstone, 2008, pp. 1-30). The term 'discourse' is central to conceptualizing this specific methodology. One of the meanings of 'discourse' is language in use or "the actual instances of communicative action in the medium of language" (ibid., p. 2). This understanding of the term enables the differentiation between language analysis and discourse analysis. The former focuses on language as an abstract system of rules and structural relationships, while the latter is interested in the way knowledge about language is used by actual communicators to perform actions. The second meaning of 'discourse', when used in the plural, is "conventional ways of talking that create and perpetuate systems of ideology, sets of beliefs about how the world works and what is natural" (ibid., p. 29). 'Analysis' within this method involves looking at the subject of the study from multiple perspectives. This results not only in descriptions of the status quo but also in critique. The main focus is on texts treated as wholes and on their contexts.

There are a number of ways for analyzing discourse. For Paltridge and Wang (2015, p. 163) the main focus of Discourse Analysis is on:

- linguistic patterns which occur across stretches of spoken and written texts;

- knowledge about language beyond the word, clause, phrase and sentence that is needed for successful communication;
- what people mean by what they say and how they work out that understanding;
- the relationship between language and the social and cultural contexts in which it is used;
- the way in which language constructs different views of the world and different understandings.

Discourse Analysis has several underlying principles as pointed out by Baxter (2010, pp. 124-125):

- *The Principle of variability*. The same phenomenon can be described in a number of different ways according to audience, purpose and context, and thus there will be considerable variation in accounts.
- *The constructed and constructive nature of language*. Discourse can never be taken as simply descriptive of the social action to which it refers, any account of experience is a form of interpretation, constituting a new version of reality.
- *The interpretative nature*. There are regular, descriptive features which are recurrently used for characterizing and evaluating actions, events and other phenomena.
- *The combination of micro- and macro-analytical approaches.* Micro-analytical approaches deal with the finer detail of language use, while macro-analytic approaches are concerned with the broader social processes that work through language. These two approaches are combined in the following way: (1) Discourse Analysis tries to identify regular features in the data such as idioms, metaphors, figures of speech and professional terminology, which may signify wider patterns of language use. (2) These in turn provide evidence for discussing the role of the psychological, social or political factors that may inform the speech or writing of research participants.

As seen from the descriptions of its key features, Discourse Analysis lacks its own formal apparatus for conducting the micro-analysis and borrows methods from other fields such as Speech Act Theory, literary criticism and Conversation Analysis (ibid., p. 126). Despite this limitation, it is a particularly effective method for studying political discourse.

4.1.2.2. Critical Discourse Analysis

Critical Discourse Analysis, also referred to as Critical Discourse Studies (Wodak and Meyer, 2016), regards itself not as a theory, a methodological approach or an explicit method, but rather as a 'critical' perspective. This is a state of mind or an attitude that can be combined with other approaches and conducted by scholars working both in linguistics and social sciences. It has developed since 1990s as a perspective that is applied by scholars from different areas who share political concerns about various social inequalities in the western world. Although multiple branches of CDA have evolved, they have one and the same linguistic starting point, i.e. Halliday's systemic functional grammar (Halliday and Matthiessen, 2014). It is underpinned by the understanding that there is a relationship between the grammatical system and the social and personal needs that language is required to serve. This relationship is maintained by the three interconnected meta-functions of language – ideational, textual and interpersonal. Therefore text and context are seen as linked in a dialectical relationship which means that they construct one another.

The underlying principles of Critical Discourse Analysis are conveniently summarized by Baxter (2010, pp. 127-128) and Ehrlich and Romaniuk (2013, pp. 477-478) in the following way:

- *Language use is seen as social practice*. This implies a dialectical relationship between a particular discursive event and the situation(s), institution(s), and social structure(s) which frame it. Thus, in this two-way relationship, discourse is considered to be socially constitutive as well as socially shaped.
- *There is a relationship between language and power*. Since discourses are so influential, they can help to produce and reproduce unequal power relations between different ethnicities, social classes, genders, ages, and professional groups.
- *Deconstruction* is the main purpose of CDA. CDA is concerned to reveal exactly how binary power relations constitute identities, subject positions and interactions within discourses and texts, and thus create social inequalities. It is believed that through analysis the demystification and denaturalizing of the opaque aspects of language is made possible.
- *An emancipatory agenda is followed*. (1) CDA focuses on major social problems, chooses the perspective of those who suffer most, and critically analyses those in power, those who are responsible, and those who have the means and the opportunity to solve such problems.

(2) CDA raises an awareness of the "naturalized" dimensions of discourse (i.e., those aspects of discourse that seem commonsensical and inevitable), with the view that such awareness may, in turn, have the effect of subverting the practices CDA analyses.

- *Text and context are closely examined.* CDA examines textual features such as sentence structure, verb tense, syntax, lexical choice, the internal coherence and cohesion of discourse, etc. This microanalysis is conducted within a 'critical perspective'. The contextual frame of the 'production' and 'consumption' of discourses is also considered. CDA involves analysis of the historical and cultural processes/structures which have given rise to the production of a text and the ways subjects within these processes/structures 'consume', or interact with texts. This implies a dialectical relationship between the reading of a particular text and the context, institution or social structure that frames this reading.
- *CDA research is characterized by self-reflexivity.* It is believed that scholarly research is never neutral and the CDA analysts should make explicit their politically engaged stance, acknowledging in a reflexive way the a priori assumptions, motivations and value systems in conducting linguistic research. Such value systems are often informed by Marxist critical theory which is considered to be offering an objective view on social reality.
- *Interdiscursivity/intertextuality is central to discourse production and comprehension.* Interdiscursivity/intertextuality involves the ways in which one discourse is always inscribed and inflected with traces of other discourses. Reader knowledge about participants, situational context and implicit power relations is relied on.

Although CDA has a broad agenda, its top-down approach has been criticized for several reasons (Baxter 2010, p. 129; Blommaert, 2005, pp. 31-37; Ehrlich and Romaniuk, 2013, p. 478). First, its methodology and analytic approaches seem rather vague. This is due to its transdisciplinary perspective and the fact that CDA practitioners have significant theoretical and methodological differences in approaching the discursive dimensions of inequality and ideology. Thus contemporary CDA is characterized by methodological pluralism. Second, its interpretations are biased, though often presented as critical. Still, CDA is acknowledged to be valuable for the study of institutional discourse because it provides models of analysis and analytical tools for deconstruction of public, media and political discourse.

4.1.3. Narrative Analysis

Narrative is often used as a synonym of 'story', 'life story', 'account', 'discourse', 'narration' and 'tale' with little or no difference in meaning. The term 'narrative' also refers to a method of analysis as in 'narrative inquiry' and is defined as any sequence of clauses which contains at least one temporal juncture (Gimenez, 2010, p. 200).

Narrative analysis or narrative inquiry is a way of doing research that focuses on the stories that are told about one's life or experiences. For Barkhuizen (2015, p. 136) stories are not merely a list of facts, but embody understandings of the events and express feelings. Experiences become narratives when told to an audience and the narratives become part of narrative inquiry when they are investigated for research purposes. Interest in narratives reflects postmodern concerns with the self, identity and individuality as they are seen as a resource that individuals draw upon in the construction of social identities. Narratives are also a particular way of constructing knowledge by organizing experience temporarily.

The study of narrative has evolved over time (Gimenez, 2010). Aristotle was the first to describe the structure of narrative plots as having a beginning, a middle and an end. Narratives have been analysed mainly following either a componential or a functional analytical approach. Componential analysis aims to identify the different elements that constitute a narrative and how these elements interact and change as a result of their interaction while the functional analysis examines the purposes of narrative. The componential approach aims to identify the basic structure of a narrative and to examine the sequence of its clauses. This sequential arrangement can then be used to determine the functions of the clauses. The other traditional way of analysing narratives is the functional approach, which mainly examines the purposes of narratives. Among the multiple functions that narrative can serve, the most widely studied is the representational function: how narrators represent or interpret the world; how they represent self and others; and how they construct their gendered, ethnic or class identities. The representational functions narratives serve can be revealed if they are studied within their macro-sociolinguistic context of production and consumption.

Scholars are paying increasing attention to the political effects of narratives. As De Fina and Johnstone (2015, p. 161) point out storytelling is "a way of creating community and as a resource for dominating others, for expressing solidarity, and for resistance and conflict". De Fina (2018) further explains how this mechanism works: (1) narratives can convey and reflect meta-discourses about society and identity; (2) through stories, both politicians and members of society at large position themselves with respect to socio-political issues; (3) narratives can be used as discursive strategies for the construction and negotiation of politically appropriate identities. In this sense, storytelling can be viewed as political practice (ibid., p. 244).

Illustrations of how narratives are used as a research method for studying political discourse can be found in:

(1) De Fina's chapter on the case of the young migrants who were taken into the US by their undocumented parents when they were children and who have since remained in the country uses the functional approach to narratives (De Fina, 2020). More specifically, the focus is on the contribution of biographical narrative to the construction of a collective identity for those youth within the context of their organized struggle for migration reform as well as for recognition of their rights and their dignity. It is argued that storytelling is a discourse practice with a central role in effort to construct and circulate a positive collective identity.

(2) Sclafani's article also uses the functional approach to narratives (Sclafani, 2015). The main focus is on candidates' self-introductions during US Republican presidential primary debates. The analysis shows how references to family members and roles, i.e. family identitity, serve as a resource to frame political identities.

4.2. Applying Quantitative and Mixed Methods

4.2.1. Questionnaire surveys

The questionnaire is the main data collection method in surveys. It has become one of the most popular research instruments applied in the social sciences and in applied linguistics in particular. The reason is that it captures the essence of scientific research by trying to find answers to questions in a systematic and disciplined way. The term 'questionnaire' is used synonymously with tests, checklists, surveys, schedules, forms, inventories, etc. However, there is an agreement among survey specialists that questionnaire is a highly structured data collection instrument. It is defined as any written instruments that present respondents with a series of questions or statements to which they are expected to react; they either write down their answers or select from existing answers (Dörnyei, 2007, pp. 101-115).

Questionnaires generate three types of data about the respondent through three different types of questions:

- Factual questions are used to find out certain facts about the respondents.
- Behavioural questions are used to find out the respondents' actions, lifestyles, habits, personal history.
- Attitudinal questions are used to find out the respondents' attitudes, opinions, beliefs, interests, and values.

The way questions are asked differentiates questionnaires from achievement or aptitude tests and from discourse completion tasks. What is special about questionnaire items is that they elicit information about the respondents in a non-evaluative manner, without assessing their performance against a set of criteria.

The wording of the items has a substantial impact on questionnaire design and on the responses. As pointed out by Rasinger "questionnaires are nothing but scientific tools that help us to measure different aspects of 'reality' […] And as such, they must measure *neutrally* and *objectively*. That means that questions should avoid as much as possible being biased or leading; in other words, they should not be phrased in such a way that they imply the 'correct' answer – quite simply because there is no 'correct' answer *per se*."(2010, p. 63).

In order to deal with the problem of the unpredictable impact of item wording multi-item scales were introduced. Thus no individual item carries an excessive load and an inconsistent response to one item would cause limited damage. Usually more than one item (the number varies between 4 to10) is used to address each identified content area, aiming to reveal slightly different aspects of it. Finally, the various multi-item scales have to be mixed up to create a sense of variety and to prevent respondents from simply repeating previous answers.

Most questionnaire items ask about specific pieces of information or give various response options. This makes questionnaire data particularly suitable for quantitative, statistical analysis but not for qualitative, exploratory research. Only rarely questionnaires comprise open-ended items that provide qualitative data. The reason from a qualitative perspective is that questionnaires direct the participants to superficial and brief engagement with the topic. Therefore most questionnaires are primarily made up of closed-ended items. The respondents are not expected to produce any writing but to choose one of the given alternatives.

The most common closed-ended item formats are the following:

- *Likert scales* consists of a characteristic statement and respondents are asked to indicate the extent to which they agree or disagree with it by marking one of the responses ranging from 'strongly agree' to 'strongly disagree'.
- *Semantic differential scales* elicit graduated response. Respondents are asked to indicate their answers by marking a continuum between two bipolar adjectives at the extremes.
- *Numerical rating scales* assign one in several numbers to describe a feature of the target. The rating continuum can refer to a wide range of adjectives and adverbs.
- Variety of *closed-ended item types* are used depending on the purpose and topic of the questionnaire, as well as on the characteristics of the respondents. These include true-false items, multiple-choice items, rank order items.

The most common open-ended item formats are *specific open questions*, *clarification questions*, *sentence completion* and *short-answer questions*. These permit greater freedom of expression and provide more richness than fully quantitative data.

The main advantage of questionnaires is their efficiency in terms of time for preparation and administering, effort for processing the data, inclusiveness of variety of respondents and topics as well as respect for ethical issues such as anonymity. The limitations of questionnaires are associated with the wording and comprehension of the items and the length and depth of the investigation.

An illustration of the application of questionnaires in the study of political discourse can be found in Kostova (2019). The paper focuses on American political discourse as represented in presidential speeches and aims to explore the way it is perceived by representatives of an outside culture. In order to study the perceptions of Bulgarian students a small-scale survey was conducted. It is based on a questionnaire which aims to gather personal, subjective views on how political text and talk is perceived and whether this influences the attitudes towards America. A combination of ten open- and closed-ended questions have been devised in line with the contemporary assumptions of American political discourse. Qualitative methods of analysis, Content and Critical Discourse Analysis in particular, have been employed to interpret and explain the results.

4.2.2. Corpus Methods

Corpus methods or corpus linguistics refer to the collection of large sets of authentic digital texts called 'corpora' and their semi-automated statistical analysis performed with computer software, i.e. 'concordance programmes'. Corpus linguistics is considered a methodology rather than a traditional branch of linguistics (Baker, 2010). Within corpus linguistics a distinction should be made between two groups of terms. First, *corpus building* means deciding what should go in the corpus while *corpus analysis* relates to what research questions should be asked, what should be looked for, what analytical procedures should be carried out, how the results can be interpreted. Second, there is a difference between *corpus-driven* and *corpus-based* approaches. In *corpus-driven* approaches a corpus is used in an inductive way in order to form hypotheses about language, not making reference to existing linguistic frameworks. In *corpus-based* approaches corpora are used in order to test or refine existing hypotheses taken from other sources (ibid., p. 95)

Corpus research originated as a distrust of both individual language biases and qualitative methods. Martinez (2020, p. 211) states that interest in enormous collections of data is associated with three issues: (1) language analysis to better understand language systems as opposed to social phenomena; (2) a way to oppose unreliable human intuition; (3) quantitative based research.

The key features of this methodology are the following as defined by Baker (2010):

1. At the centre of corpus linguistics is the concept of the corpus. Any text or collection of texts could be theoretically conceived of being a corpus. However, a corpus normally consists of a sample that is maximally representative of the variety under examination is of a finite size, exists in machine readable form, and constitutes a standard reference for the language variety which it represents. This means that it will be large enough to reveal something about frequencies of certain linguistic phenomena, enabling researchers to examine what is typical, as well as what is rare in language.
2. Corpus size is dictated by a number of criteria. The most significant one concerns the aspects of language that the corpus is used to investigate. For example a corpus of 100,000 words will usually be big enough to make generalizations for most descriptive purposes; studying other aspects of language use would require half a million or a million words.
3. Sampling, balance and representativeness are key theoretical concepts in corpus linguistics. Because a corpus ought to be representative of a particular language, language variety, or

topic, the texts within it must be chosen and balanced carefully in order to ensure that some texts do not distort the corpus as a whole. So corpora may not contain whole texts, instead utilizing parts of texts.

4. Corpora are often annotated (or tagged) with additional information, allowing more complex calculations to be performed on them. It is also possible to tag a corpus for other types of linguistic information.
5. A range of different types of corpora are in existence. First, a distinction needs to be made between general and specialized corpora. A *general corpus* is one which aims to be representative of a particular language, such as the British National Corpus, The Bank of English or the American National Corpus. A *specialized corpus*, however, can be smaller and contains a more restricted set of texts in terms of genre, time, place/language variety. Specialized corpora are generally easier than general corpora to collect and are used to answer specific research questions. However, specialized corpora are often used in conjunction with *general corpora*, with the general corpus acting as a 'benchmark' about typical language, being compared to the specialized corpus in order to show what forms of language (e.g. lexis, grammar, topics) are over- or underrepresented in the smaller corpus. Another distinction involves whether a corpus contains spoken, written or computer-mediated texts (such as emails, text messages or websites) or a mixture of all three. A third distinction involves the language or languages which a corpus is encoded in. A *multilingual corpus* usually contains equal amounts of texts from a number of different languages, often in the same genre. A *parallel corpus* is a more carefully designed type of multilingual corpus, where the texts are exact equivalents (i.e. translations) of each other. A *learner corpus* is a corpus of a particular language produced by learners of that language.
6. Corpora are normally used in conjunction with analysis software, which are able to carry out the counting, sorting and presentation of language features, such as WordSmith Tools, Xaira, Wmatrix and AntConc.
7. Many forms of corpus-based analysis are based around the concept of *frequency*. The most basic aspect of frequency analysis simply allows us to derive frequencies of particular words (or phrases or tags), or lists of all of the words in a corpus, presented alphabetically or in order of frequency. Presenting frequencies as percentages is often useful, particularly when making comparisons between multiple corpora, especially of different sizes.

8. A related form of frequency analysis involves calculating *keywords*. A keyword is a word which occurs statistically more frequently in one file or corpus, when compared against another comparable or reference corpus. Some keywords can be explained due to events or people who were particularly in vogue at the time when the corpora were collected. Other words suggest more subtle social changes or are indicative of changes in style.
9. It is often not enough to simply extrapolate explanations based on the presence of keywords alone. They need to be investigated in more detail and in context. So how can we investigate context? This is where the concept of the concordance is useful. A concordance is simply a list of a word or phrase, with a few words of context either side of it, so we can see at a glance how the word tends to be used. Corpus analysis software normally allows concordances to be sorted alphabetically in various ways (e.g. one, two, three, etc. words to the left or right of the word under examination), which allows humans to recognize patterns more easily. The examination of concordances also helps to reveal discourse prosodies. Discourse prosodies are often indicative of attitudes. A concordance analysis therefore combines aspects of quantitative and qualitative analyses together.
10. Because corpora can contain thousands or millions of words, this can often result in a huge amount of information to analyse by hand. A statistical procedure which helps to reduce this information to more manageable chunks is *collocation*. Collocation refers to the statistically significant co-occurrence of words. There are a number of different ways of calculating collocation. Some, like the mutual information score (i.e. which take into account exclusivity of collocation – for example, words must always appear together and not apart), tend to give precedence to low frequency collocations involving nouns, adjectives and verbs. Other ways of calculating collocation, such as log likelihood, which gives precedence to highly frequent collocates, tend to favour grammatical relationships.

As corpus linguistics is mainly characterized as a methodology, it can be used in a number of different ways:

- It can aid both linguistic description and language teaching.
- It can add systematicity to and reduce subjectivity in stylistic analysis.
- Humans do not always make accurate introspective judgements regarding language, instead relying on cognitive and social biases. Computers can calculate frequencies and

carry out statistical tests quickly and accurately, giving researchers access to linguistic patterns and trends – such as collocational information that were previously inaccessible.

- It can enable researchers to confirm or refute hypotheses about language use, as well as allowing them to raise new questions and theories about language that otherwise would not have been possible.
- It can enable researchers to quantify linguistic patterns, providing more persuasive conclusions to be reached – large corpora allow researchers to find evidence of rare or unusual cases of language, as well as shedding light on very frequent phenomena.

Corpus analysis is of additional value to the study of social and political issues for several reasons (Kutter, 2018, p. 183):

- It provides a distinct view of texts by the abstraction that it creates from large amounts of texts.
- It does not present texts according to their linear or intertextual composition. Instead, it highlights patterns of regular use of single words or word clusters as they appear in a large collection of texts. Such reading provides a bird's eye view of the use of specific words *across* all the texts contained in a corpus.
- The display of word clusters reveals patterned semantic, syntactic, and other, relationships between words. This gives an additional layer of interpretation, which cannot be gained otherwise.
- It may challenge not only the habitual way of linear reading, but also the preconceived views about the subject studied.

However, corpus linguistics is not able to answer every research question in the area of linguistics. The following limitations are significant (ibid., p. 184; Baker, 2010):

- The results of corpus exploration are often not meaningful in themselves. In order to relate to discourse studies of social and political issues, they have to be made part of an analytical strategy.
- Corpus analysis does not help in reconstructing the larger context of utterance.
- Corpus analysis selectively focuses on the distributional properties of words. Therefore it is not suited to achieve deep comprehension of texts and their contextual and

interdiscursive conditioning. This remains in the domain of interpretive or detailed linguistic discourse analysis.

- Corpora can be time-consuming, expensive and difficult to build, requiring careful decisions to be made regarding sampling and representativeness.

Consequently, it is suggested that corpus analysis can be applied in the study of political discourse only as an explorative technique. Mixed with other qualitative methods of textual analysis, it can give insights into texts and facilitate generating hypotheses and research questions.

4.2.3. Mixed Methods

A useful methodological synergy is combining corpus linguistics and Critical Discourse Analysis (Mautner, 2016). This involves the following steps:

- Compiling an electronically held corpus that allows the investigation of research questions arising from social issues.
- Running the corpus through concordancing software that compiles frequency lists, identifies keywords and reveals statistically significant collocations.
- Analyzing concordances qualitatively in order to establish the dominant semantic preferences and prosodies of lexical items relevant to the social issues under investigation.
- Putting the results from purpose-built corpus into perspective by comparing them with evidence from large reference corpora.
- A multi-million word reference corpus may allow researchers to build collocational profiles of contested lexical items across a wide range of genres, media and geographical areas.

There are clear benefits in combining the two methods: (i) corpus linguistics allows researchers to closely examine much larger and potentially more representative amounts of textual data; (ii) this strengthens the empirical foundations of the studies, reduces researchers' bias and enhances the credibility of analysis; (iii) the integration of the two methods is a significant move towards methodological pluralism; (iv) corpus linguistics is also enriched by being applied to research questions inspired by social concern, such as power, inequality and change.

Examples of combining quantitative and corpus methods with Critical Discourse Analysis to examine political discourse can be found in the following research articles:

(1) Degani (2016) studied populism realized through anti-intellectualism in American political discourse from a linguistic perspective by comparing communication strategies in the announcement speeches of two major American presidential candidates. Anti-intellectualism is seen both as related to linguistic complexity as measured by certain textual features and as connected to characteristics of discourse such as representation of Self, representation of Other, and framing of political issues. This integrated approach is based on both quantitative and qualitative methods of analysis.

The purely quantitative analysis of linguistic complexity of the political speeches as a measure of their ant-intellectualism is carried out by eight readability tests for estimating the difficulty of written language. Based on mathematical formulas, they rely on either the syntactic factor (number of sentences, number of words per sentence, number of syllables per word and number of characters) or a combination of a syntactic factor with a semantic one (difficulty of words). The results of these statistical measures have shown that there is a significant difference between the two speeches in terms of text simplicity.

A combination of corpus and qualitative methods have been used to examine framing of discourse. First, the thirty most frequent words in the speeches as well as the frequency of occurrence of each word have been considered. Second, the listed words have been manually extracted from two distinct Wordlists generated by Wordsmith tools after excluding all instances of function words. The selected categories of content words includes nouns, adjectives, verbs (excluding modals and auxiliaries) and adverbs with adjectival base. Finally, the analysis of the idiosyncratic lexical choices of the selected classes of words, including the use of personal pronouns, relies on the concepts and analytic instruments of Systemic Functional Linguistics. Thus the discourse analysis of the high frequency lexical items and their socio-political context of use confirms the major differences between the messages constructed by the two speakers.

(2) Rivers & Ross (2020) examined authority legitimation and delegitimation in the border wall Twitter discourse of President Trump. Two research methods were used consecutively – corpus approaches and Critical Discourse Analysis. The research process encompassed the following procedures. First, Twitter posts were scraped from the personal Twitter account of Trump using the FireAnt application. Second, collected Twitter posts were subject to a corpus analytical approach of a comparative keyword analysis. The research corpus, i.e. the collected

Tweets, were compared to a discourse reference corpus, i.e. Political Twitter Discourse Corpus. Third, a comparative keyword analysis was undertaken using AntConc application which generates a keyword list comprised of words that are unusually frequent within the research corpus. The keyword analysis was carried out for the purposes of sampling. The key words 'border' and 'wall' were isolated as the basis for more focused qualitative investigation because quantitative corpus-based analysis in isolation reveals nothing about the strategies of legitimation and delegitimation within the Tweets. Finally, qualitative analysis was conducted by focusing specifically on those tweets that included the most frequent keywords.

Conclusion

Although this paper does not claim to provide a complete theoretical overview of the general research approaches, the main types of research, the key research methodology issues and the specific research methods in linguistics, it has thrown light on important principles and techniques for conducting research. It has shown that the relation between *method*, *methodology* and the *epistemological aspect* of research paradigm is circular. On the one hand a specific paradigm, either realistic or constructivist, explains the nature of knowledge and has an impact on methodology. On the other hand specific methodology suggests and justifies specific methods that collect and analyse data. The analysis conducted under a specific method or a combination of methods leads to interpretation, i.e. finding the answers to the research questions, which creates knowledge. Thus the primary issue that the paper attempted to discuss is how various research approaches transform their theoretical concepts and assumptions into procedures, instruments and methods of analysis, i.e. how theory is operationalized. A conclusion can be drawn that in the research process, theory, methods and analysis are closely interrelated and decisions made about one affect the others.

The research methods of data collection and analysis within contemporary linguistics presented in the paper are justified mainly by social constructionism but they differ in many respects. First, they rely on various linguistic theories. Second, they integrate linguistic categories into their analysis to a different extent and with a different focus and intensity. This means that they keep different balance between macro- and micro-analysis. The most evident similarity between some of them is their interest beyond language as they aim to shed light on the discursive aspects of some social processes. In relation to studying political discourse there is a clear trend towards applying mixed methods and triangulation to meet the quality criteria of research.

References

Angouri, J. (2010). Quantitative, qualitative or both? Combining methods in linguistic research. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 29-47). London/ New York: Continuum.

Baker, P. (2010). Corpus methods in linguistics. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 93-115). London/ New York: Continuum.

Barkhuizen, G. (2015). Narrative inquiry. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: a practical resource* (pp. 136-148). London/New York: Bloomsbury Academic.

Baxter, J. (2010). Discourse-analytic approaches to text and talk. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 117-137). London/ New York: Continuum.

Benson, P. (2013). Qualitative methods: overview. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 5188-5295). Malden/Oxford/Chichester: Wiley-Blackwell.

Blommaert, J. (2005). *Discourse. A critical introduction*. Cambridge: Cambridge University Press.

Creese, A. (2010). Linguistic ethnography. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 138-154). London/ New York: Continuum.

De Fina, A. (2018). Narrative analysis. In R. Wodak & B. Forchtner (Eds.), *The Routledge handbook of language and* politics (pp. 233-246). London/New York: Routledge.

De Fina, A. (2020). Biography as political tool: The case of the dreamers. In *Sociolinguistic Perspectives on Migration Control* (pp. 64-85). Multilingual Matters.

De Fina, A. & Johnstone, A. (2015). Discourse analysis and narrative. In D. Tannen, H. Hamilton & D. Schiffrin (Eds.), *The handbook of discourse analysis* (2nd ed.) (pp.152-167). London: Wiley-Blackwell.

Degani, M. (2016). Endangered intellect: A case study of Clinton vs Trump campaign discourse. *Iperstoria – Testi Litterature Linguaggi* 8: 131-145.

Dörnyei, Z. (2007). *Research methods in applied linguistics. Quantitative, qualitative and mixed methodologies.* Oxford: Oxford University Press.

Duranti, A. (2006). Narrating the political self in a campaign for U.S. Congress. *Language in Society* 35(4): 467-497.

Ehrlich, S. & Romaniuk, T. (2013). Discourse analysis. In R. Podesva and D. Sharma (Eds.), *Research methods in linguistics* (pp. 460-493). Cambridge: Cambridge University Press.

Farhady, H. (2013). Quantitative methods. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 5196-5209). Malden/Oxford/Chichester: Wiley-Blackwell.

Gimenez, J. (2010). Narrative analysis in linguistic research. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 198-216). London/ New York: Continuum.

Halliday, M. A. K. & Matthiessen C. (2014). *Halliday's introduction to functional* grammar (4th ed.). London/New York: Routledge.

Hepburn, B. & Andersen, H. (2021). Scientific method. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Available at: https://plato.stanford.edu/archives/sum2021/entries/scientific-method/

Holliday, A. (2015). Qualitative research and analysis. In B. Paltridge & A. Phakiti (Eds.). *Research methods in applied linguistics: A practical resource* (pp. 49-59). London/New York: Bloomsbury Academic.

Ivankova, N. & Greer, J. (2015). Mixed methods research and analysis. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource* (pp. 60-74). London/New York: Bloomsbury Academic.

Kostova, B. (2019). American political discourse: outsider perceptions and attitudes. *Studies in Linguistics, Culture and FLT, 5,* 31-45.

Krzyżanowski, M. (2018). Ethnography and critical discourse studies. In J. Flowerdew & J. Richardson (eds.), *The Routledge handbook of critical discourse studies* (pp. 179-194). London/New York: Routledge.

Levon, E. (2013). Ethnography and recording interaction. In R. Podesva and D. Sharma (Eds.), *Research methods in linguistics* (pp. 195-215). Cambridge: Cambridge University Press.

Litosseliti, L. (Ed.). (2010). *Research methods in linguistics*. London/ New York: Continuum.

Mautner, G. (2016). Checks and balances: how corpus linguistics can contribute to CDA. In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse analysis* (3rd ed.) (pp. 154-179). London: Sage.

McKinley, J. (2020). Introduction: Theorizing research methods in the 'golden age' of applied linguistics research. In J. McKinley and H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 1-13). New York: Routledge.

Parmelee, J. & Perkins, S. (2013). Discourse – mixed methods. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1966-1971). Malden/Oxford/Chichester: Wiley-Blackwell.

Phakiti, A. & Paltridge, B. (2015). Approaches and methods in applied linguistics research. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource* (pp. 19-33). London/New York: Bloomsbury Academic.

Phakiti, A., De Costa, P., Plonsky, L. & Starfield, S. (2018). *The Palgrave handbook of applied linguistics research methodology*. London: Palgrave Macmillan.

Paltridge, B. & Wang, W. (2015). Discourse analysis. In B. Paltridge and A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource* (pp. 162-174). London/New York: Bloomsbury Academic.

Rasinger, S. (2010). Quantitative methods: concepts, frameworks, and issues. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 49-67). London/ New York: Continuum.

Reisigl, M. & Wodak, R. (2016). The discourse-historical approach (DHA). In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse analysis* (3rd ed.) (pp. 23-61). London: Sage.

Rivers, D. & Ross, A. (2020). Authority (de)legitimation in the border wall Twitter discourse of President Trump. *Journal of Language and Politics* 19(5): 831-856.

Sclafani, J. (2015). Family as a framing resource for political identity construction: Introduction sequences in presidential primary debates. *Language in Society* 44(3): 369–399

Starfield, S. (2015). Ethnographic research. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource* (pp. 112-135). London/New York: Bloomsbury Academic.

Sunderland, J. (2010). Research questions in linguistics. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 9-28). London/ New York: Continuum.

Wodak, R. (2012). Politics as usual: Investigating political discourse in action. In P. Gee & M. Handford (eds.), *The Routledge handbook of discourse analysis* (pp. 525-540). London/New York: Routledge.

Wodak, R. & Forchtner, B. (Eds.). (2018). *The Routledge handbook of language and politics.* London/New York: Routledge.

Wodak, R. & Meyer, M. (2016). Critical discourse studies: history, agenda, theory, and methodology. In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse analysis* (3rd ed.) (pp. 1-22). London: Sage.