VNU Journal of Foreign Studies

Journal homepage: https://jfs.ulis.vnu.edu.vn/

METHODS FOR AMASSING DATA IN CROSS-CULTURAL PRAGMATICS RESEARCH

Le Huu Loc*

Department of Foreign Languages, Quy Nhon University,
170 An Duong Vuong, Quy Nhon, Binh Dinh, Vietnam

Received 18 February 2024
Revised 29 July 2024; Accepted 19 November 2024

Abstract: This article discusses prevalent and up-to-date data-amassing techniques for cross-cultural pragmatics studies. Based on the distinctive features of data types, the commonly used methods are outlined with their pluses and minuses, along with the author's remarks and recommendations on how to employ them for fulfilling research purposes. Additionally, the paper suggests that an integration of the data-gathering methods proves to be of great value for pragmatics studies in general and cross-cultural pragmatics studies in particular, contributing to ensuring the validity and reliability of these kinds of research.

Keywords: cross-cultural research, data, data collection, methodology, pragmatics

* Corresponding author.
Email address: lehuuloc@qnu.edu.vn

https://doi.org/10.63023/2525-2445/jfs.ulis.5235

PHƯƠNG PHÁP THU THẬP DỮ LIỆU CHO NGHIÊN CỨU DỤNG HỌC GIAO VĂN HÓA

Lê Hữu Lộc

Khoa Ngoại ngữ, Trường Đại học Quy Nhơn,
170 An Dương Vương, Quy Nhơn, Bình Định, Việt Nam

Nhận bài ngày 18 tháng 02 năm 2024
Chỉnh sửa ngày 29 tháng 7 năm 2024; Chấp nhận đăng ngày 19 tháng 11 năm 2024

Tóm tắt: Bài viết này thảo luận các kỹ thuật thu thập dữ liệu phổ biến và cập nhật cho các nghiên cứu dụng học giao văn hóa. Dựa trên đặc điểm của các loại dữ liệu dùng cho nghiên cứu dụng học giao văn hóa, bài viết cung cấp một cái nhìn tổng quan về các phương pháp thu thập thường được sử dụng, đi kèm với phân tích về những ưu và nhược điểm của từng phương pháp. Từ đó, bài viết đưa ra những nhận xét và khuyến nghị về cách sử dụng các phương pháp nhằm đạt được mục đích nghiên cứu. Đồng thời, bài viết cũng chỉ ra giá trị của việc kết hợp các phương pháp thu thập dữ liệu cho nghiên cứu ngữ dụng học nói chung và dụng học giao văn hóa nói riêng nhằm góp phần đảm bảo tính tin cậy và chính xác cho những nghiên cứu này.

Từ khóa: nghiên cứu giao văn hóa, dữ liệu, thu thập dữ liệu, phương pháp, ngữ dụng học

1. Introduction

House and Kádár (2022) define “cross-cultural pragmatics” as “the field of studying language use in a contrastive way across languages and language varieties.” (2022, p. 151). Cross-cultural pragmatics is distinctively characterized by its cross-disciplinary relevance as it has been employed in studies of applied linguistics, linguistic politeness, and translation. Therefore, the results of these studies feature a high level of implication in various areas, especially language teaching and learning.

Data can be considered of crucial importance in cross-cultural pragmatics studies since the selection of appropriate, reliable, and compatible data sources plays a huge part in ensuring the research’s validity and reliability. The justification for method choice, however, seems not to receive sufficient awareness in conducting research, leaving the procedure of data-gathering normally ignored or unelaborated.

This paper is intended to outline data-amassing methods in cross-cultural pragmatics based on the nature of a variety of data types. Besides, the strengths and weaknesses of each method are discussed with recommendations offered regarding how to take advantage and maximize the benefits and minimize its drawbacks, fulfilling particular aims of research.

2. Methods in Cross-Cultural Pragmatics Research

2.1. Classification of Data-Amassing Methods in Pragmatics

Schneider (2018) groups the data collection methods in pragmatics into three categories, including “institution”, “observation”, and “experimentation”. In a similar fashion, these three categories also coincide with the three metaphors, “armchair”, “field”, and “laboratory” as proposed by Clark and Bangerter (2004, p. 25), which are employed by Jucker (2009) in his review of methods for carrying out studies on the speech act of compliment.

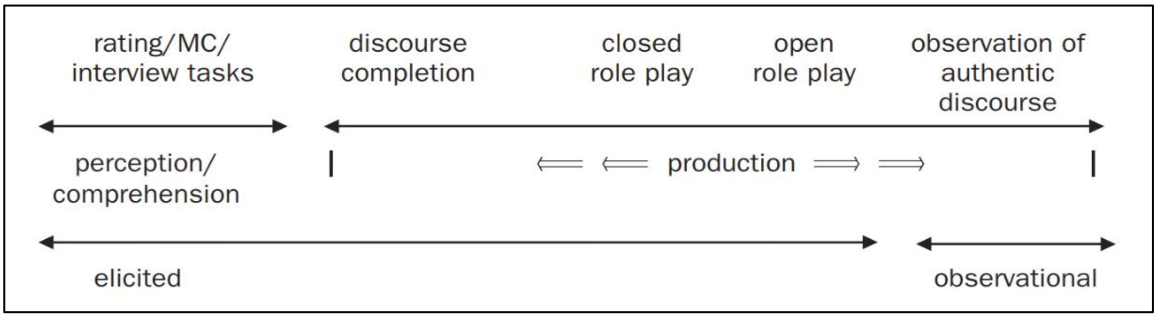
More concretely, preliminary armchair research involves researchers entirely basing on their experience in communication and competence in pragmatics. This is also considered "second-order introspection" (Schneider, 2018, p. 49), and it proves of great use in forming theories and proposing communication principles and maxims. Indeed, the method of armchairing has been put into use by linguists to develop such theories as speech act theory (Austin, 1962), relevance theory (Sperber & Wilson, 1996), politeness theory (Brown & Levinson, 1987; Lakoff, 1973; Leech, 2014; Leech, 1983), Co-operative Principle (CP) (Grice, 1975) and Politeness Principle (PP) (Leech, 2014; Leech, 1983), conversational maxims (Grice, 1975) and politeness maxims (Leech, 2014; Leech, 1983). "Armchair method" is not considered an empirical method since the researchers' intuition is not data regarding the sense in which this concept is frequently employed, with no instruments and specific lines of actions adopted by the researchers. On the other hand, "field" methods concern the use of empirical tools of investigation to make analyses of authentic natural language uses as discovered "out in the field", whereas researchers using "laboratory" methods concentrate on conducting experiments in a laboratory for data elicitation from appropriate respondents.

Tracing the origins of the three methods, Hodeib (2021) clarifies that armchair methods have their roots in corpus and descriptive linguistics. In detail, in formal linguistics, armchair linguists rely on their institutions to make evaluations on the standard of language use, whereas they allude to the use of corpora in the context of computer meditation as regards corpus linguistics (Fillmore, 1992). Laboratory linguists make use of the methods from psycholinguistics and experimental linguistics in order to gather data in controlled settings with varied methods of data elicitation such as roleplay. Regarding field methods, which are originally from the domain of sociolinguistics, field linguists embrace the observation of natural authentic language, and bring into service a wide range of data resources namely emails and text messages.

Kasper and Dahl (1991) suggest that research methods in pragmatics can be orderly represented on a scale as in Figure 1. "Elicited" and "observational" are placed at the two extremes of the scale. Each sort of data requires distinct collecting instruments to be employed. To illustrate, rating tasks and multiple-choice questions are the typical tools for obtaining elicited data that concern respondents' awareness and comprehension of the particular pragmatic phenomenon examined, whereas collecting naturally occurring data pertains to observing authentic discourse.

Figure 1

Data Collection Methods, Represented as a Continuum Between Elicitation and Production

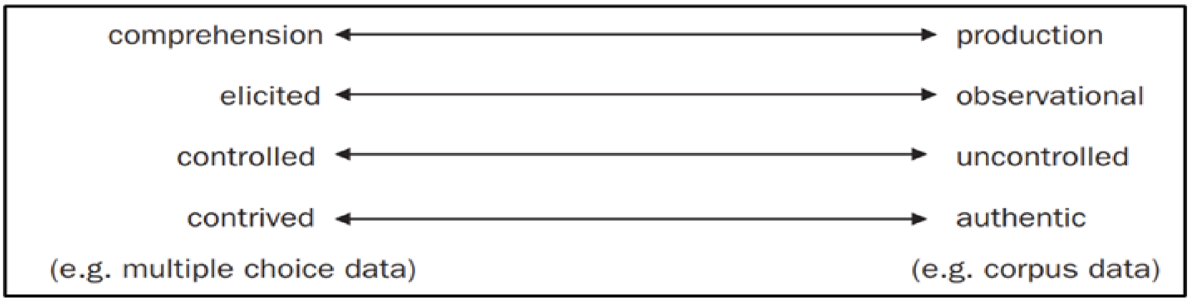


(Kasper & Dahl, 1991; Leech, 2014)

Elaborating on the continuum as proposed by Kasper and Dahl (1991), Leech (2014) outlines four ways of portraying the extremes of a methodological continuum as presented in Figure 2.

Figure 2

Four Ways of Characterizing Poles of Methodological Continuum (Leech, 2014)



As regards the pair of “controlled” and “uncontrolled” methods, multiple-choice tests and corpora are the two typical types of data-gathering instruments. In detail, while a multiple choice test “constrains” a participant to decide upon one of several possible responses or utterances, leaving out the possibility for other responses, the corpora encompass conversational data gathered with no specified constraints on what the participants are permitted to say to each other. As a result, proof from corpora characterizes the uncontrolled pole of the continuum.

Concerning the extremes of “contrived” and “authentic”, corpus data are naturally occurring in real life, whereas the responses in a multiple choice task are considered contrived or unreal since they are grounded on what the respondent thinks or imagine he/she may say in envisaged settings. A multiple-choice questionnaire can be viewed from the perspective of a source for proof of metapragmatic awareness, and this can be mined from the metapragmatic task which requires the participants to pass judgment on the pragmatic properness of the utterances given. Leech (2014) asserts this type of multiple-choice task “is far removed from the real language-use data that pragmatics in principle should investigate” (2014, p. 249)

2.2. Types of Data in Pragmatics Research

According to Jucker (2018), “the data of any pragmatic research is the language used in actual contexts, and language is ever pervasive” (2018, p. 4). He further clarifies that “pragmaticians” (the linguists specializing in pragmatics) have directly or indirectly displayed a preference for particular kinds of data instead of other types, putting much emphasis on unconstrained spoken interaction in authentic environments. On the other hand, written language has frequently been disregarded as inadequate for pragmatic analysis due to its secondary nature. Besides, fictional language has encountered more unwillingness to be employed as a result of its artificiality. Nonetheless, even particular kinds of spoken language created in the classroom or courtroom are highly constrained because, in such circumstances, the allocation of roles and legitimate contributions of the communicators are explicitly made.

Apart from the distinction of spoken versus written language data, according to Jucker (2018), there are other ways of conceptualizing the kinds of data of pragmatics research, namely, online/ digital data, sign language data, and data of nonverbal behavior. He also proposes four dimensions on which the taxonomy of data, particularly observational data, can be established, including situational dimension (constrained versus unconstrained data), fictionality dimension (fictional versus factual data), researcher interference (low interference versus high interference data, as represented in Figure 3) and research perspective (micro versus macro data, as displayed in Figure 4).

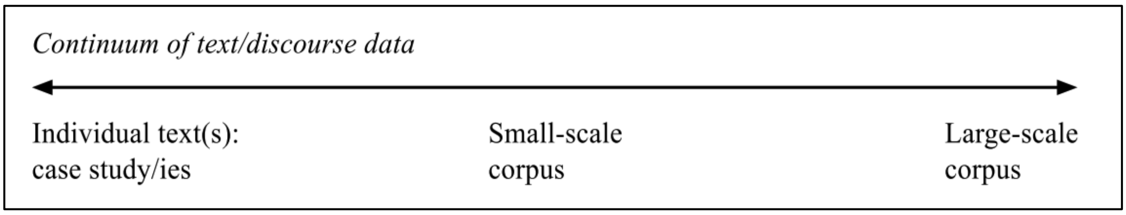
Figure 1

Data Collection as Regards Researcher Interference (Jucker, 2018)

Researcher Interference	Control		Relevant examples
Low	Low	1	Speech recording without researcher involvement
		2	Surreptitious recording by researcher
		3	Non-surreptitious recording by researcher
		4	Participant observation recording
		5	Semi-structured interview
		6	Role play or role enactment
		7	Dialog construction task
High	High	8	Oral DCTs

Figure 2

Data Collection as Regards Researcher Perspective (Jucker, 2018)



The dimension of the researcher's perspective concerns the size of the data examined. As can be seen in Figure 4, at one polar of the scale, a very small amount of normally amply contextualized data is employed, typically a single dialogue or even only a small excerpt of a dialogue with the information regarding the interactants and the setting of the dialogue is well informed. At the other extreme of the scale, the research needs to resort to the use of large corpora of millions or billions of words in search of language use patterns from those corpora.

Bednarek (2011) outlines two categories of data in pragmatics research, namely "attested data" and "non-attested data". She defines "non-attested data" as follows:

> any kind of data that has not been observed or recorded as having been produced by a particular speaker or writer. This covers predominantly the use of introspection - thinking and reasoning about language use based on one's own knowledge and experience of language. (2011, p. 539)

Introspection is deemed to play a vital role in developing theories, and forming hypotheses; however, these theories or hypotheses need to be applied to attested data so that they can be modified, confirmed, or rejected. In fact, it would be contended by many researchers of corpus linguistics that there is no certain correspondence between what is thought about language use and the actual language use. Phrased another way, introspection seems to concentrate more on the possibilities in language use, rather than on the authentic, significant language use.

On the other hand, Bednarek (2011) defines "attested data" as follows:

> Attested data means language data which was actually produced by speakers/ writers on particular occasions. Such data can be elicited (by the researcher) or can have occurred more "naturally", that is, without being produced in a research context. (2011, p. 540)

Attested data is comprised of "elicited data" and "non-elicited data". More concretely, elicited data involves interviews (concerning the researcher), verbal reports, elicited narratives, discourse completion tasks, and experimental data, whereas non-elicited data encompasses observational ethnographic data and diverse types of text/discourse data, including corpora.

Regarding the dichotomy of elicited versus non-elicited data, House and Kádár (2021, p. 43) claim "there has been a strong disagreement between various strands of research favoring one element of the naturally occurring versus elicited binary pair over the other". To illustrate, the Cross-Cultural Speech Act Realisation Project (CCSARP) (Blum-Kulka et al., 1989) is criticized as the data employed in this project is regarded as elicited and "decontextualized". Even yet, the CCSARP Project found great value in the elicited and decontextualized data since it made it possible to gather, at the time, previously unheard-of levels of pragmatic data, which in turn allowed for a highly quantitative analysis of pragmatic phenomena. Therefore, this kind of data proves to be beneficial to cross-cultural pragmatics research that aims at reaching cross-cultural results that can be replicated and generalized. On the other hand, the non-elicited data seems to be more suitable for "discursive research" and qualitative scrutinization which concern the researchers relying on a small amount of naturally occurring cross-cultural data and are likely to verge on peculiarity rather than replication.

Kirner-Ludwig (2022), in a review of data collection methods applied in the studies in the journal *Intercultural Pragmatics,* summarizes three kinds of data in relation to how they are garnered, namely "observed"/ "observational"/ "naturally-occurring", "extracted", and "elicited". In detail, naturally occurring data is gathered "through external, non-biasing observation of subjects and communicative scenarios" (2022, p. 461). As Jucker (2009) explains, this kind of data "has not been elicited by the researcher for the purpose of his or her research project but that occurs for communicative reasons outside of the research project for which it is used" (2009, p. 1615), which is the principle underlying the field method. However, collecting "observational" data encounters the problems of lack of constraint of the speaker and variables of settings, participants' age, and their social backgrounds. Furthermore, securing the validity and reliability of this sort of data entails taking numerous factors into account, such as whether the recording is taken properly. Another data type that is often used in field methods is "extracted" data, which is characterized as "quantifiable, large sets of materials (available as electronic corpora)" or "individual texts". While observed data is, on the whole, naturally occurring and genuine, extracted data is dynamic on a scale of "naturalness". Schneider (2018) justifies,

> corpus data do not all qualify as observational data. They are naturally occurring to the extent that their existence does not depend on a researcher. Yet there are significant differences between the data types included in machine-readable corpora, sometimes even in the same corpus. (2018, p. 50)

The third type of data as outlined by Kirner-Ludwig (2022) is "elicited" data. This data is obtained through "applying systematically controlled settings, prompts, and variables" (2022, p. 462). Schneider (2018) puts forward a flexible grouping of eliciting tasks as regards low, medium, or high levels of interaction/ collaboration needed. More concretely, he states that tasks of elicitation should be considered on a "continuum [...] decreasing [in] interactionality and, at the same time, increasing [with regard to] researcher control" (2018, p. 58). Low-interaction tasks involve very limited interaction extent to elicit productive data. Examples of low-interaction tasks are closed-ended questionnaires (intuitive tasks to ask for respondents' age, gender, occupation, etc.), diary/verbal reports (gathering participants' relevant anecdotes or self-observations), open-ended questionnaires, or written/oral DCT (Discourse Completion Task). Medium-interaction tasks are typically elicited dialogues in which "researchers specify topics, interactional goals or discourse roles" (Kasper, 2008, p. 287). Interviews are salient examples of medium-interaction tasks, which can be divided into, full versus in-part narrative (unstructured, open-ended), semi-structured versus structured, formal versus informal, and audio- versus video-recorded. Finally, high-interaction tasks are connected with "collaborative learning activities of various kinds employed to elicit productive data" (Kirner-Ludwig, 2022, p. 464). Illustrations of high-interaction data can be found in collaborative writing or translation assignments, video-conferencing sessions, group discussions, peer feedback tasks, and role-play tasks (RPTs).

2.3. Empirical Methods for Amassing Data in Cross-Cultural Pragmatics

Based on examining the observational data in the preceding part, this section serves to dwell on outlining the common experimental data-amassing methods in cross-cultural pragmatics, including "applying a corpus", "recording authentic talk", "delivering production tasks", and "delivering comprehension and judgment tasks" (Schneider, 2018). Each method is presented with the primary contents of its core components coupled with its pluses and minuses as well as suggestions of how it can be appropriately applied according to the researcher's purposes.

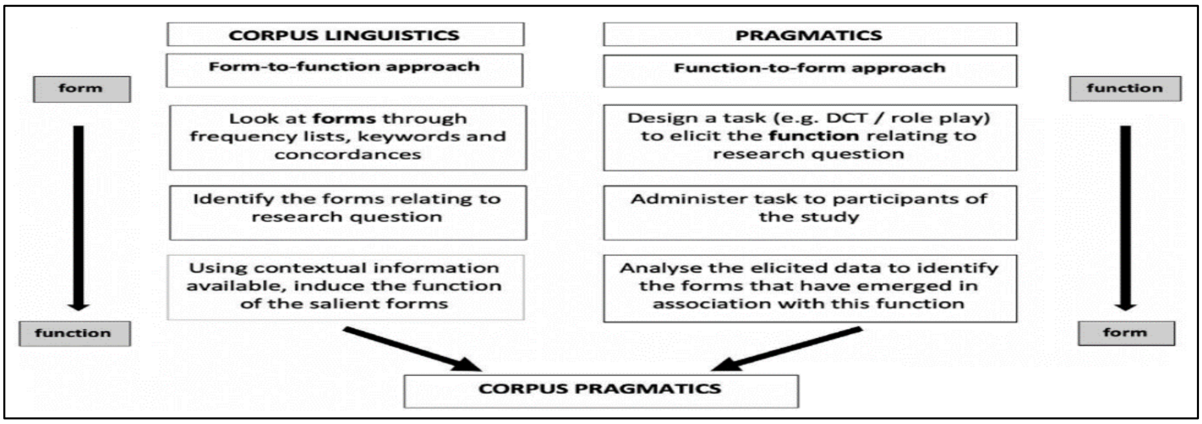
2.3.1. Applying a Corpus

Unlike armchair pragmatics, which merely relies on the researchers' intuitive knowledge to form hypotheses or develop theories, empirical pragmatics requires the use of data corpora. This holds true for both fieldwork and laboratory work, with corpora of observational data and corpora of experimentational data, respectively (Schneider, 2018, p. 50). A corpus can be understood as a very large electronic collection of texts in written or spoken form, compiled not for any research purposes. The British National Corpus (BNC), the Corpus of Contemporary American English (COCA), and the national or regional corpora belonging to the International Corpus of English (ICE) are typical examples of corpora. In a broader sense, a corpus can be any collection of texts, no matter how large it is or whether or not it can be read by machines. Therefore, a corpus can be a self-compiled one, which is much smaller than the huge corpora and is adjusted to meet particular research purposes.

Although Kasper and Dahl (1991) do not mention using corpora in the continuum in Figure 1, it is possible to contend that "electronic corpora, as they have developed over the past twenty-five years, provide the most readily available instrument for "observing authentic discourse"." (Leech, 2014, p. 256). As Schneider (2018) states, applying a corpus proves useful when researching micro units, namely discourse markers and similar concepts since form-based corpus can provide researchers with great assistance in promptly and exhaustively searching for such micro units. According to O'Keeffe et al. (2020), while pragmatics has conventionally worked from function to form, corpus linguistics works in the opposite direction, starting from form; corpus pragmatics is the use of corpus linguistics in pragmatics, with the combination of the form-to-function and function-to-form approach. They clarify that "[t]he challenge for the emerging model of CP [corpus pragmatics] research is to avoid assuming that form-to-function processes are the only option and to find ways of addressing the challenges to function-to-form approaches (2020, p. 48). This is characterized through Figure 5.

Figure 3

Approaches in Discourse Pragmatics (O'Keeffe et al., 2020)



Corpora can be applied to studies of particular speech acts. For these studies, illocutionary force indicating devices (IFIDs), such as performative verbs (e.g. invite, offer, apologize) and other devices are normally employed in realizing the speech acts of corresponding speech acts, namely inviting, offering, and apologizing. However, as O'Keeffe et al. (2020) point out, "[a]n IFID can prove unreliable as a means of recalling all, and only, instances of a speech act" (2020, p. 53). To deal with this issue, it is recommended that researchers search both speech act verbs (e. g. invite, suggest, warn) and speech act nouns (e. g. invitation, suggestion, warning) for speech acts' realizations and speech acts discussion (e. g. reporting, commenting, challenging) (Jucker, 2012).

The obstacles to using corpora for cross-cultural research concern the comparability of corpus data across corpora; the appropriateness of corpus data in comparative studies remains restricted since corpus data are not spontaneously comparable. Another difficulty facing studies employing corpus data is the shortage of information regarding participants' backgrounds, and this is genuinely a problem for studies, especially those of variational pragmatics, targeting the effects of macro-social factors, including respondents' age, gender, or region.

2.3.2. Recording Authentic Talk

Data-gathering techniques are also as broad as the wide range of communication technology. Indeed, as for recording authentic talk, the levels of complexity of the recording technique depend on the particular kinds of data. For instance, while text-based computer-mediated interaction data (via emails, discussion groups, chat forms, etc.) can be recorded in a rather straightforward manner, voice or visual access entails recording technology and arrangements being more complex (Markham, 2004). According to Kasper (2008), regarding the study of speech acts, there are three kinds of recording techniques, including "field note", "audio recording" and "video recording" (2008, p. 284). All three kinds of recording techniques can vary in terms of size, scope, and fineness of what is recorded; however, they all adhere to the principle of being "perspectival", "selective" and "partial" (2008, p. 284)

It is widely acknowledged that the consent must be approved by respondents before recording, and the "observer's paradox" emerges, as justified by Labov (1973):

> [...] the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain this data by systematic observation. (1973, p. 209)

Phrased another way, the talk cannot be recorded without participants being alert, and as a result, they act in a less natural manner. Indeed, as Schneider (2018) points out, "[f]or researchers, "audio- or video-recording naturally occurring conversation in the truest sense of the word is virtually impossible" (2018, p. 53). In the same vein, House and Kádár (2021) warn that "fetishising 'naturally occurring' is a grave error because such data may never be genuinely naturally occurring – this entirely depends on how we define the concept of 'naturally occurring-ness'" (2021, p. 49). However, it is informed by Schneider (2018) that, in some studies, the longer the speech event and recording proceed, the more naturally participants behave and feel relaxed, especially when they are with friends and in familiar surroundings.

One strategy that is commonly used by researchers of conversational analysis or interactional linguistics is concentrating on recording a particular type of discourse and/or a specific kind of setting (Schneider, 2018, p. 54). This can also be related to the technique employed in discourse analysis, in which the researcher records "stretches of discourse, sometimes quite extensive stretches" (Leech, 2014, p. 256), to investigate the discursive aspect of politeness, for instance. By doing this, the researchers are enabled to detect the patterns of interaction examined that are normally left unnoticed. However, the downsides are that few excerpts of discourse can be thoroughly studied, and the selection of excerpts is frequently influenced by the analyst's individual interests. However, the pervasive use of video recording allows for preserving discourse data in visual and sound forms or in transcribed forms, which can greatly assist the researcher in scrutinizing authentic speech.

Whether the analyst participates in the interaction is also an issue that is worth considering. If the researcher plays the role of a participant-observer, it provides him/her with the opportunity to gear the direction of the interaction towards the aims and research questions of the study. Nevertheless, this can diminish the level of naturalness and originality of the interaction. Otherwise, an external observer will not have the opportunity to gain access to pertinent information and will be likely to pass misjudgments regarding the participants' relationship, particularly in daily exchanges. If the researcher is not well informed of who the interactants are and what they have exchanged prior to observation/ recording, it is likely that the researcher fail to comprehend what is communicated. One way to deal with this situation can be to supplement the recording with a follow-up interview with the participants where the transcripts can be discussed so that the researcher can yield a more sufficient comprehension of the interaction and the researched phenomenon.

Another technique for collecting authentic data is "taking field notes" (Kasper, 2008, p. 284) or "the notebook method" (Jucker, 2009, p. 1616). This is a typical technique in ethnographic research, which primarily involves overhearing what other people say and noting it down by hand. One benefit of this approach is that it does not require the consent of those being overheard, thereby avoiding the observer's paradox. Additionally, there is no transcribing work required and no electronic recording equipment required (Schneider, 2018, p. 55). This method has been popular in sociolinguistic research, and thus appropriate to be adopted in speech act research. For instance, Holmes (1986) employed field notes to scrutinize compliments and responses to compliments in New Zealand English. The limitations of applying this instrument include the fact that the researchers focus on the prominent and explicit speech act realizing strategies, leaving more implicit and authentic strategies unregistered. Another shortcoming of taking field notes results from inaccurate hearing and memorizing. Kasper (2008) clarifies, "what can be recorded by researcher's observation and subsequent field notes is constrained by human cognitive capacities". Because attention is selective and memory information degrades swiftly, people's short-term memories are ineffective as a documenting tool for interactions. Consequently, only single-turn, brief, high-frequency "semantic formulae" including greetings, leave-takings, and (certain types of) compliments can be reliably seen through conventional ethnographic investigation. On the other hand, the chronological and prosodic arrangement of a speech act, as well as its consecutive structure, which describes how the main action is prompted, how it is reacted to, and how it evolves over possibly many turns, loses shape in memory. (2008, p. 285).

Furthermore, researchers employing the technique of recording authentic talk may encounter difficulties in accessing the research site over a sufficient period of time for audio- or video- recording. In addition, they may find it hard to gather enough comparable data from participants with diverse features, including languages or language proficiency. Moreover, the practice of transcribing demands a huge amount of time, experience, and even specialized training. The transcription practice in question involves the measurement of pauses, interruptions, overlaps, and concurrent talk, not to mention the transcription of non-verbal behavior.

2.3.3. Delivering Production Tasks

Production tasks refer to the kinds of methods to gather data by having participants produce language. The thread among these methods is that they are experimental methods. As such, they meet the criteria summarized by Schneider (2018) as follows

> (1) The language produced does not occur naturally, i.e. it does not arise from the genuine needs and desires of language users, but occurs on the initiative of a researcher. (2) The language produced is elicited under conditions determined by the researcher, sometimes referred to as "laboratory conditions". That is to say, the researcher usually decides on the time, place, and setting of the data elicitation. (3) All language users serving as informants are consciously aware that they are involved in an experiment and that their language productions are recorded, not necessarily electronically, and then used for research purposes. To this, they have given their consent, and they participate voluntarily. (4) All informants follow instructions and complete a task designed by the researcher. (5) At least in most cases, the language produced does not have any social consequences, unlike naturally occurring discourse. This lack of consequences contributes to the often bemoaned artificiality of the elicitation situations and the language produced therein. (2018, pp. 57-58)

Elicited conversation, role plays, interviews, and discourse completion tasks are among the production tasks covered in this area. They can be viewed as constituting a continuum with growing researcher control and dwindling interactionality.

2.3.3.1. Elicited Conversation

Elicited conversation engages two participants, who are not asked to take any social roles but essentially be themselves. The topic of conversation is not determined in advance. Kasper (2008, p. 287) considers elicited conversation as tasks whose topics, communicative purposes, and discourse roles are specified by the researchers. Nonetheless, those tasks are synonymous with role-play or role enactment.

Elicited conversation exhibits all prosodic, formal, actional, interactional, organizational, etc. aspects of spontaneously occurring conversation that are pertinent to pragmatics research. As a result, collections of this data type may be used as corpora for researching a variety of phenomena, such as intonation, discourse markers, speech act realization, adjacency pairs, speech act sequences, conversational openings and closings, turn-taking, interruptions, and silence.

Although elicited conversations can provide insight into a variety of interactional resources and practices, some study objectives necessitate that researchers exercise experimental control over participant roles, contextual factors, and communicative activities.

2.3.3.2. Roleplays

Role plays can be generally understood as mimics of communication encounters that are typically (but not always) carried out in pairs according to role descriptions or guidelines. Role-plays are characterized as “a social or human activity in which participants “take on” and “act out” specified “roles, often within a predefined social framework or situational blueprint (a “scenario”)” (Crookall & Saunders, 1989, pp. 15-16).

Félix-Brasdefer (2018, p. 308) summarizes two major kinds of role plays, closed and open role plays. Closed roleplays aim to generate one-turn replies and are fundamentally oral discourse completion tasks. The oral response to the question is captured on tape. These are suggested to elicit longer responses than written DCTs and to obtain verbal and non-verbal characteristics of spoken discourse that a written DCT will not, such as hesitation, backchannels, gestures, etc. An open role-play, on the other hand, is characterized by being two-way. Participants are typically instructed to role-play how they would reply after reading the scenario prompt. The conversation is taped for perusal at a later stage.

O’Keeffe et al. (2020) assert that role plays are frequently employed for data elicitation in cross-cultural and interlanguage pragmatics, particularly in making comparisons between native (NS) and non-native speaker (NNS) language use in relation to the same task (NS - NS versus NNS - NNS) or within the same task (NSs-NNSs). (2020, pp. 30-31)

Boxer and Cohen (2004, p. 17) claim that under certain circumstances, role-playing data are similar to naturally occurring spoken data, provided that the researcher could provide a setting in which conversation is examined. Demeter (2007), who uses role play to collect data for research on apologies, believes that role play can offer researchers the opportunity to get closer to actual data when examining the speech acts’ production. This is confirmed by Rosendale (1989), who employs role plays to elicit data for a study on the speech act of invitation.

Despite being criticized for artificiality, role plays prove a useful instrument in investigating interactional angles of communication (including turn organization, turn taking and overlap, etc.), as well as prosodic characteristics and hints (namely intonations, tone, and stress). Additionally, they can make it possible to investigate macro-social aspects (such as gender, age, ethnicity, and socioeconomic position), both within and between languages.

2.3.3.3. Interviews

Interviews are common in all facets of social life, and since they are such a well-known source of knowledge, they are the most widely used method for gathering qualitative data. (Dörnyei, 2007; Rolland et al., 2020). In applied linguistics and pragmatics research, interviews are highly prevalent, and they are used to capture a wealth of information, including the participants’ attitudes about a certain issue and their self-reflections regarding language-related behavior. (Rolland et al., 2020). Kasper (2008) argues that interviews can be implemented as the primary data collection instrument or as a triangulation approach. (2008, p. 296). However, she also advises against treating interviews as precise externalizations of consistent, decontextualized thoughts and knowledge levels; interview responses are co-constructed by nature since interviews are structured as question-answer chains, and interviewees’ replies are therefore always impacted by the questions. (2008, p. 296)

According to Dörnyei (2007, pp. 135-136), based on the researchers' objectives, there are three types of interviews to be selected, namely "structured", "unstructured" and "semi-structured" interviews. Researchers often employ structured interviews to elicit concentrated information on a particular area of interest. They are favored when the researcher is aware of their limitations and is able to use inquiries to uncover crucial information. Contrarily, unstructured interviews give the respondents discretion over the conversation's course in an effort to elicit personal narratives. The most common interview type is the semi-structured interview, which allows the interviewee to go off course and produce meaningful material in unexpected places or ways while still following a predetermined set of questions.

Kasper (2008) claims that interviews can serve diverse research purposes. First, researchers can ask participants in meta-pragmatic interviews about their first-order conceptualizations of speech acts or categories including politeness and rudeness. These interviews can be used to compare second-order conceptualizations established in armchair research with first-order conceptualizations (lay people's interpretations), with the goal of improving the completeness of theoretical constructs and definitions. Interviewees may also be prompted to give instances of a specific speech act, such as a complaint or an invitation, for the purpose of data collection. These interviews resemble closed role plays or oral discourse completion tests. Additionally, interviewees can give instances of circumstances in which small talk is probable, compliments are expected, or examples of what they would say in a specific situation. This can be used to create scenarios for discourse completion tasks or role plays, improving their ecological validity. Finally, post-production interviews/retrospective interviews can be used to fully comprehend participants' selections, realizations, and decisions.

In pragmatics research, interviews are a common technique that is frequently combined with other methods of data collecting to examine phenomena. For instance, Blum-Kulka (1992) examined Israeli attitudes and beliefs toward politeness using structured, semi-structured, and open interviews. Japanese and American participants' displays of attentiveness are examined in Fukushima's (2011) cross-cultural study. Both questionnaires and interviews were used to gather the data, with the interview data revealing detailed insights into the participants' decisions and corroborating the questionnaire results. Fukushima (2013) focuses on the numerous aspects determining impoliteness and looks at how attentively participants from Japan and America evaluate it. These issues are addressed thanks to the combination of survey and interview data. Besides, the perceptions of (im)politeness of a naturally occurring apology in Australian English and Taiwanese Mandarin are investigated by Chang and Haugh (2011) and Haugh and Chang (2019) through interviews and rating scales. The follow-up interviews helped to clarify the moral standards and cultural schemata that underpinned the participants' judgments of (im)politeness and the intercultural and intracultural diversity in perceptions that are evident in the respondents' rating choices.

However, interviews should also be conducted with care. There is always a potential that the social desirability bias will impact responses if the interview questions are not stated neutrally enough. Additionally, the participants can be either too shy to provide enough commentary or overly talkative, which would result in lengthy but unusable data. (Dörnyei, 2007, p. 144). Schneider (2018) adds that in the follow-up interviews of role plays, comments and explanations can be elicited by having the participants read the transcripts or watch the video recordings; however, as a result of the delay, respondents may no longer remember some facts with accuracy or be able to tell the interviewer what they believe they said and why, which can result in inaccurate data. (2018, p. 64)

2.3.3.4. Discourse Completion Tasks

In cross-cultural and interlanguage pragmatics, Discourse Completion Tasks (DCTs) have been a widely utilized and harshly criticized elicitation method. The format was initially created by Levenston and Blum (1978) to examine lexical simplification, and then it was modified by Blum-Kulka (1982) to look at speech act realization. Kasper and Dahl (1991) define DCTs as tasks that "are written questionnaires including a number of brief situational descriptions, followed by a short dialogue with an empty slot for the speech act under study" (1991, p. 221). For this task, participants are requested to provide a response that they believe matches the scenario. As Kasper (1999) explains, it is possible to further restrict the participant's written response by allowing a rejoinder to come after the available slot.

As Ogiermann (2018) points out, DCTs are of great value for cross-cultural and interlanguage pragmatics studies. She justifies that the DCT seems to be the only data collection tool that can produce sufficiently large corpora of comparable, systematically diverse speech act data, which is necessary for research that aims to identify culture-specific patterns in speech act realization or the pragmatic characteristics of a particular interlanguage (2018, p. 229). DCTs are a useful solution for the contrastive research of speech acts as they can be converted into any language and swiftly delivered to large respondent groups. (Aston, 1995; Barron, 2003, p. 85). Taking the same stance, Schneider (2018) asserts that written DCTs offer the benefit of enabling the gathering of language samples from a large number of respondents in a short amount of time, for instance by delivering a written DCT to a large lecture hall, or by email, social media, or crowdsourcing platforms. Another benefit is that extensive DCT data collection could provide the researcher with valuable metadata for correlation work (e.g. across gender, nationality, first language, age, etc.).

The Cross-Cultural Speech Act Realization Project (CCSARP), carried out by a multinational team of linguists, is the largest and most significant DCT study to date (Blum-Kulka et al., 1989). The experiment analyzed requests and apologies in five languages (Canadian French, Danish, German, Hebrew, and English), with the latter one represented by three variations (American, Australian, and British). The CCSARP paradigm has been used recurrently in speech act investigations, yielding a substantial corpus of comparable data from a wide range of additional languages. Numerous DCT studies have concentrated on requests and/or apologies and closely followed the project's design. This is made easier by duplicating the CCSARP DCT, or a modified version of it, and by the fact that the two speech acts generated for the project have a thorough coding scheme available. (Ogiermann, 2018, p. 230)

Multiple researches comparing the validity of DCT data to other elicited data and to naturally occurring discourse have been conducted in the past two decades (Economidou-Kogetsidis, 2013; Golato, 2003; Yuan, 2001). The findings of these research, which are generally similar, indicate that DCT data are far more suited to inquiries into how much respondents know about what is suitable to say in specific settings as opposed to answering questions about how language is realized spontaneously. Sweeney and Hua (2016, p. 215) summarize the discrepancies between DCT data and other kinds of elicited data and authentic speech acts. More concretely, according to Golato (2003), despite certain parallels between the strategies as found in DCTs and those in the naturally occurring data, other DCT strategies do not appear in the natural data that was recorded, and vice versa. Besides, in DCTs, responses are provided to prompts where in natural data they would not be. Additionally, while replies in elicited data are delivered over one turn, responses in the natural data emerge over multiple turns (Golato, 2003; Yuan, 2001). Moreover, responses to DCTs may not be comparable since they depend on the design of the questionnaire (Billmyer & Varghese, 2000). As regards the types of DCTs, the responses to oral DCTs are more in line with the natural data than written DCTs (Yuan, 2001).

As DCT requires participants to "recall pragmatic information from memory and report rather than use it" (Barron, 2003, p. 85), one of the main criticisms of DCTs has been that the responses do not often match what the speakers would say if they were in the scenarios that are being described, but rather what they believe they would say (Aston, 1995, p. 62; Schneider, 2011, p. 18). However, considering that the goal of cross-cultural pragmatic studies is to demonstrate general, culture-specific patterns of language usage, the potential mismatch between the responses and actual language use does not inherently invalidate DCT findings. It is not important whether the participants would use the same language if they were placed in the scenarios presented, as long as they believe their responses to be socially and culturally acceptable (Ogiermann, 2018, p. 233).

2.3.4. Delivering Comprehension and Judgment Tasks

Generally, studies that use comprehension and judgment tasks focus mostly on individual utterances rather than interactional discourse. As a result, such tasks are typically not appropriate for macro-pragmatic analysis, such as the investigation of discursive sequences or whole speech events. This section discusses multiple choice questions and rating scales, two of the main categories of comprehension and judgment tasks.

2.3.4.1. Multiple Choice Tasks

According to Kasper (2008), multiple-choice (MC)/multiple-choice tasks (MCTs) have been employed for investigating "people's preferences for speech act strategies and forms, comprehension and metapragmatic judgments" (2008, p. 294). Similar to DCTs, MC items provide a prompt for a response and define the situational background, but instead of requiring the participant to actively create a response, MC gives a number of alternative responses from which one must be opted.

Schneider (2018) states that "MCTs seem particularly suitable to investigations into impoliteness and verbal aggression, e.g. swearing or insults, as informants are often inhibited when they are requested to commit "foul language" to paper" (2018, p. 70). He adds that respondents tend to prefer MCTs to DCTs due to the lower processing demands involving MCTs regarding the cost of cognition than the recalling tasks like DCTs.

However, the reliability of MCTs is also challenged in recent studies concerning speech act production. (Brown, 2001; Röver, 2005; Yamashita, 1996). The unreliability seems to point to the reality that many speech acts can be performed in a multitude of context-appropriate ways, which is a premise of pragmatic life. In addition, unless the research focuses on a very restricted object, identifying potentially suitable combinations of strategies of realization and linguistic forms in a principled manner is a challenging issue for instrument designers.

According to Röver (2005), MCs that examine respondents' comprehension of speech acts and implicature and that require the choice of contextual patterns can be developed with sufficient levels of reliability.

2.3.4.2. Rating Scales

Rating scales are regarded as a subtype of the questionnaire by Kasper (1999). Leech (2014) asserts that they are a version of multiple-choice tasks. In any case, rating scales require participants to convey their assessment or perception of a particular pragmatic phenomenon by selecting a particular value or label from a variety of alternatives. Participants may also be asked to reorder a group of linguistic expressions in accordance with a particular standard of evaluation, such as appropriateness, politeness, or grammaticality. The use of rating scales is common not only in pragmatics study but also in sociolinguistics research and theoretical linguistics research, particularly experimental syntax, to elicit grammaticality judgments. In general, the purpose of applying rating scales, particularly the Likert scale is to "elicit assessments of utterances or situations in terms of correctness, appropriateness, politeness, formality, and so on" (Schneider, 2018, p. 70). Scales with a typical range of 1 to 5 are used for this purpose. (Dörnyei, 2003, pp. 36-39).

Rating scales are employed in cross-cultural and interlanguage pragmatics to generate pragmalinguistic and sociopragmatic evaluations. In the former, participants are asked to assess whether linguistic realizations of a specific speech act are suitable in a particular social setting, while in the latter, participants are questioned about how they interpret and comprehend a situation with regard to aspects such as power, distance, degree of imposition, and severity of offense. The construction of tasks like role plays and discourse completion can greatly benefit from sociopragmatic assessments. In a pre-test, participants may be asked to score the contextual factors included in the scenarios that researchers have sketched out for such tasks. This practice contributes to raising the validity of the study (Kasper, 2008, pp. 295-296)

Rating scales in politeness research are less common now than they were before discursive techniques emerged, which argue that politeness cannot be assessed in isolation and outside of context (Watts, 2010). However, Leech (2014, pp. 250-251) believes that participants' judgments of what he refers to as "pragmalinguistic politeness" represent a default perception and that it can be evaluated outside of context, for example, by ranking alternate realizations of a speech act. Here, it is important to emphasize that these default judgments display first-order conceptualizations of politeness. In addition, Leech disagrees with Kasper and Dahl's (1991, p. 219) critique that, in a judgment exercise, participants must create their own situational context if the researcher fails to do so. He contends that there is no evidence to support their premise and asserts that participants rely on a "generalized context". Think-aloud protocols that are simultaneously recorded while judgment tasks are being completed could provide some evidence.

Rating scales prove helpful in swiftly gathering assessments from a large number of respondents and assigning quantitative values to qualitative data for statistical analysis (Collins et al., 2009). However, choosing the proper kind of scale is contingent on the quantity of values and labeling. As Schütze (2016) argues, too many value points can elicit "spurious distinctions" among participants, while too few can obscure true distinctions. Short scales seem less demanding and more reliable for statistical analyses. Additionally, labeling values on scales should take into account participants' inability to make judgments regarding targeted structure, as labeling "not sure" may leave out valid cases. (Ellis, 1991, as cited in Schütze (2016)).

3. Conclusion

Starting from the categorization of types of data in pragmatics research, the paper delves deeper into the classification of data-gathering methods for pragmatics studies in general and cross-cultural studies in particular. Indeed, particular data types concern specific data-collecting instruments chosen, which also justifies the occasional overlaps between data types and data-amassing methods.

The author focuses on the observational data with its four dimensions, laying the foundation for reviewing the major types of empirical data-gathering methods, including "applying a corpus", "recording authentic talk", "delivering production tasks", and "delivering comprehension and judgment tasks". These can be seen as the most commonly adopted methods in cross-cultural and interlanguage pragmatics. Besides, other data-amassing methods for pragmatics research, namely the "philological method", "diary method", "verbal reports"/ "self-reports", or "think-aloud protocol", and methods inherited from psycholinguistics and neurolinguistics are out of scope of the current paper.

It is widely acknowledged that studies that pragmatics research in general and cross-cultural pragmatics studies in particular aim at examining language in interaction or a particular context. Therefore, they are multifaceted with a diverse range of data types and data-gathering instruments developed. However, it would be inappropriate to consider that there is only one particular method. All the methods introduced in this paper are feasible ways of collecting empirical data in cross-cultural pragmatics studies, with distinct benefits and drawbacks; the researchers are expected to be open-minded as regards employing more than one approach to create possible triangulation and to increase the validity and reliability of their research. Indeed, cross-cultural pragmatics research commonly calls for "a multi-method approach" or "componential approach" (House & Kádár, 2021, p. 44) featuring the integration of "contrastive" and "ancillary" research. In detail, the ancillary methods, such as utilizing interviews, questionnaires, and translations, can serve in a pilot study prior to the fundamental contrastive research or as a way of verifying the validity of the contrastive analysis (2021, pp. 3–4). Furthermore, it is advisable for the researcher to establish connections between the narrow and detailed cases and the more expansive and generalized hypotheses (Chafe, 1994, p. 10). Similarly, gathering and analyzing data in cross-cultural pragmatics studies involves expanding upon the results of specific cases to generate conclusions regarding wider cross-cultural patterns. After the prevailing cross-cultural patterns are uncovered, studies on particular cases can be carried out for the purpose of validity affirmation. Last but not least, on account of the diverse data types and data-amassing methods, careful consideration should also be taken to ensure the compatibility of the contrastive data sources to yield results regarding similarities and discrepancies of linguistic features and speech acts with a high level of accuracy and persuasiveness.

References

Aston, G. (1995). Say 'Thank you': Some pragmatic constraints in conversational closings. *Applied Linguistics, 16*(1), 57–86.

Austin, J. L. (1962). *How to Do Things with Words*. Oxford University Press.

Barron, A. (2003). *Acquisition in interlanguage pragmatics: Learning how to do things with words in a study abroad context (Vol. 108)*. John Benjamins Publishing.

Bednarek, M. (2011). Approaching the data of pragmatics. In W. Bublitz & N. R. Norrick (Eds.), *Foundations of Pragmatics* (pp. 537–560). De Gruyter Mouton. https://doi.org/10.1515/9783110214260.537

Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests. *Applied Linguistics, 21*(4), 517-522. https://doi.org/10.1093/applin/21.4.517

Blum-Kulka, S. (1982). Learning how to say what you mean in a second language: A cross-cultural study of Hebrew and English. *Applied Linguistics, 3*(1), 29-59.

Blum-Kulka, S. (1992). The metapragmatics of politeness in Israeli society. In R. J. Watts, S. Ide, & K. Ehlich (Eds.), *Politeness in Language: Studies in its History, Theory and Practice* (pp. 255-280). De Gruyter Mouton. https://doi.org/doi:10.1515/9783110886542-013

Blum-Kulka, S., House, J., & Kasper, G. (1989). Investigating Cross-Cultural Pragmatics: An Introductory Overview. In *Cross-cultural pragmatics: Requests and apologies* (Vol. 31, pp. 1-34).

Boxer, D., & Cohen, A. D. (2004). *Studying Speaking to Inform Second Language Learning (Vol. 8)*. Multilingual Matters.

Brown, J. D. (2001). Pragmatics tests: Different purposes, different tests. In G. Kasper & K. R. Rose (Eds.), *Pragmatics in Language Teaching* (pp. 301–326). Cambridge University Press. https://doi.org/10.1017/CBO9781139524797.020

Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge University Press.

Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing.* Chicago & London: University of Chicago Press.

Chang, W.-L. M., & Haugh, M. (2011). Evaluations of im/politeness of an intercultural apology. *Intercultural Pragmatics, 8*(3), 411–442. https://doi.org/10.1515/iprg.2011.019

Clark, H. H., & Bangerter, A. (2004). Changing Ideas about Reference. In I. A. Noveck & D. Sperber (Eds.), *Experimental Pragmatics* (pp. 25–49). Palgrave Macmillan UK. https://doi.org/10.1057/9780230524125_2

Collins, C., Guitard, S., & Wood, J. (2009). Imposters: An online survey of grammaticality judgments. *NYU Working Papers in Linguistics, 2*, 1–22.

Crookall, D., & Saunders, D. (1989). *Communication and simulation: From two fields to one theme* (Vol. 4). Multilingual matters.

Demeter, G. (2007). Symposium article: Role-plays as a data collection method for research on apology speech acts. *Simulation & Gaming, 38*(1), 83–90.

Dörnyei, Z. (2003). *Questionnaires in second language research*. Lawrence Erlbaum.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies.* Oxford University Press.

Economidou-Kogetsidis, M. (2013). Strategies, modification and perspective in native speakers' requests: A comparison of WDCT and naturally occurring requests. *Journal of Pragmatics, 53*, 21–38.

Félix-Brasdefer, J. C. (2018). Role plays. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in Pragmatics* (Vol. 10, pp. 305–331). Walter de Gruyter.

Fillmore, C. J. (1992). "Corpus linguistics" or "Computer-aided armchair linguistics". In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991* (pp. 35-60). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110867275.35

Fukushima, S. (2011). A cross-generational and cross-cultural study on demonstration of attentiveness. *Pragmatics, 21*(4), 549-571. https://doi.org/10.1075/prag.21.4.03fuk

Fukushima, S. (2013). Evaluation of (im)politeness: A comparative study among Japanese students, Japanese parents and American students on evaluation of attentiveness. *Pragmatics, 23*(2), 275-299. https://doi.org/10.1075/prag.23.2.04fuk

Golato, A. (2003). Studying Compliment Responses: A Comparison of DCTs and Recordings of Naturally Occurring Talk. *Applied Linguistics, 24*(1), 90-121. https://doi.org/10.1093/applin/24.1.90

Grice, H. Paul. (1975). Logic and Conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Academic Press. http://www.communicationcache.com/uploads/1/0/8/8/10887248/logic_and_conversation.pdf

Haugh, M., & Chang, W.-L. M. (2019). "The apology seemed (in) sincere": Variability in perceptions of (im) politeness. *Journal of Pragmatics, 142*, 207-222.

Hodeib, C. (2021). An overview of research methods in speech acts and politeness. The case for triangulation. *ARGUMENTUM, 17*, 108–124. https://doi.org/10.34103/argumentum/2021/7

Holmes, J. (1986). Compliments and compliment responses in New Zealand English. *Anthropological Linguistics*, 485–508.

House, J., & Kádár, D. Z. (2021). *Cross-Cultural Pragmatics.* Cambridge University Press. https://doi.org/10.1017/9781108954587.006

House, J., & Kádár, D. Z. (2022). Research Report: Cross-cultural Pragmatics. *Acta Universitatis Sapientiae, Philologica, 14*(2), 151-156. https://doi.org/10.2478/ausp-2022-0021

Jucker, A. H. (2009). Speech act research between armchair, field and laboratory: The case of compliments. *Journal of Pragmatics, 41*(8), 1611–1635. https://doi.org/10.1016/J.PRAGMA.2009.02.004

Jucker, A. H. (2012). Pragmatics in the history of linguistic thought. In K. Allen & K. M. Jaszczolt (Eds.), *The Cambridge Handbook of Pragmatics* (pp. 495–512). Cambridge University Press.

Jucker, A. H. (2018). Data in Pragmatic Research. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in Pragmatics* (pp. 3–36). De Gruyter Mouton.

Kasper, G. (1999). Data collection in pragmatics research. *University of Hawaii Working Papers in ESL, 18*(1), 71-107. http://hdl.handle.net/10125/40802

Kasper, G. (2008). Data Collection in Pragmatics Research. In H. Spencer-Oatey (Ed.), *Culturally Speaking: Culture, Communication and Politeness Theory* (pp. 279–303). Continuum.

Kasper, G., & Dahl, M. (1991). Research Methods in Interlanguage Pragmatics. *Studies in Second Language Acquisition, 13*(2), 215–247. https://doi.org/10.1017/S0272263100009955

Kirner-Ludwig, M. (2022). Data collection methods applied in studies in the journal Intercultural Pragmatics (2004-2020): A scientometric survey and mixed corpus study. *Intercultural Pragmatics, 19*(4), 459–487. https://doi.org/10.1515/ip-2022-4002

Labov, W. (1973). *Sociolinguistic Patterns* (*Issue 4*). University of Pennsylvania Press.

Lakoff, R. T. (1973). The logic of politeness; or, minding your p's and q's. In C. Corum, T. Cedric Smith-Stark, & A. Weiser (Eds.), *Papers from the 9th Regional Meeting of the Chicago Linguistic Society* (pp. 292-305). Chicago Linguistic Society.

Leech, G. (2014). *The Pragmatics of Politeness.* Oxford University Press.

Leech, G. N. (1983). *Principles of Pragmatics.* Longman.

Levenston, E. A., & Blum, S. (1978). Discourse-Completion as a Technique for Studying Lexical Features of Interlanguage. *Working Papers on Bilingualism, 15*, 13–21.

Markham, A. N. (2004). Internet Communication as a Tool for Qualitative Research. In Silverman D. (Ed.), *Qualitative Research: Theory, Method and Practice* (pp. 95–124). Sage.

Ogiermann, E. (2018). Discourse Completion Tasks. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in Pragmatics* (pp. 229–256). De Gruyter Mouton.

O'Keeffe, A., Clancy, B., & Adolphs, S. (2020). *Introducing Pragmatics in Use* (2nd ed.). Routledge.

Rolland, L., Dewaele, J.-M., & Costa, B. (2020). Planning and Conducting Ethical Interviews: Power, Language and Emotions. In *The Routledge Handbook of Research Methods in Applied Linguistics* (pp. 279-289). Routledge.

Rosendale, D. (1989). Role-play as a data-generation method. *Simulation & Games, 20*(4), 487-492.

Röver, C. (2005). *Testing ESL Pragmatics*: *Development and Validation of a Web-Based Assessment Battery*. Peter Lang Verlag.

Schneider, K. P. (2011). Imagining conversation: How people think people do things with words. *Sociolinguistic Studies, 5*(1), 15-36. https://doi.org/10.1558/sols.v5i1.15

Schneider, K. P. (2018). Methods and ethics of data collection. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in Pragmatics* (pp. 37-94). De Gruyter Mouton. https://api.semanticscholar.org/CorpusID:64874874

Schütze, C. T. (2016). *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Language Science Press.

Sperber, D., & Wilson, D. (1996). *Relevance: Communication and Cognition* (2nd edition). Blackwell.

Sweeney, E., & Hua, Z. (2016). Discourse Completion Tasks. In Z. Hua (Ed.), *Research Methods in Intercultural Communication: A Practical Guide* (pp. 212-232). Wiley Blackwell.

Watts, R. J. (2010). Linguistic politeness theory and its aftermath: Recent research trails. In M. Locher & S. Graham (Eds.), *Interpersonal Pragmatics* (pp. 43-70). https://doi.org/10.1515/9783110214338.1.43

Yamashita, S. O. (1996). *Six measures of JSL pragmatics.* University of Hawaii Press.

Yuan, Y. (2001). An inquiry into empirical pragmatics data-gathering methods: Written DCTs, oral DCTs, field notes, and natural conversations. *Journal of Pragmatics, 33*(2), 271-292. https://doi.org/10.1016/S0378-2166(00)00031-X