CORPUSES APPROACHES TO DISCOURSE ANALYSIS

1.   **Central Idea of the Chapter**   Discourse analysis is basically a research activity i.e. a micro level work on the text / spoken **discourse as a single whole conveying meanings** collectively. In corpus analysis **the data is fed to computers and modern gadgetry** to get **certain out puts based on the information given in the discourse**. Corpus studies can make **important contribution** to our understanding of **data / characteristic** of discourse. However, the completely computer based analysis can make **unnecessary generalizations which can lead to unrealistic outcomes and judgments**.

2.   **What is Corpus**Corpus is the collection of data usually stored and analyzed electronically. They look at **the occurrence and re occurrence** of particular linguistic features to see how and where they occur in the discourse.

3.   **Kinds of Corpora.**There are following **eight kinds** of corpora:-
   a.   General Corpora.
   b.   Specialized Corpora. .
   c.   The Michigan Corpus of Academic Spoken English.
   d.   The British Academic Spoken English.
   e.   The British Academic Written English Corpus.
   f.   The TOEFL Spoken and Written Academic Language Corpus

4.   **Design and Construction of Corpora**.   Data for general research is always **available in abundance**. However if answer to a **specific questions** is required then we have to **design our own corpora** to suit our requirements.

5.   **Issues to consider in constructing a Corpus**. While designing a Corpus many factors need to be considered, like the **size of corpus, form of expression like written or spoken genre or in the form of monologue, dialogic and multi-party**etc.
   a.   **Authenticity, representativeness and Validity of the Corpus**. In the construction ofcorpus **authenticity, representativeness and validity are also important issues**, that need due attention of the researchers.

b. **Kinds of Texts to Include in the Corpus**. Accordingto **(Hunston2002)**, the**kind of Corpus depends on the availability of texts** as well as the validity.He means to say that, for how much time it will be available for thepurpose of updating.

c. **Size of the Texts in the Corpus**. **It depends** that, some corpora aim for an even **sample size of an individual texts**.

d. **Sampling and representativeness of the Corpus**. The important issue in this regard to define **what section of the society is being considered** as **the population** being researched on.

6. **The Longman Spoken and Written English Corpus**. **The LSWE (Longman Sspoken and Written English)** corpus represents four major discourse types, **conversation, fiction, news and academic prose**.

7. **Discourse Characteristics of Conversational English**

Following are the major characteristics as noted by Longman Grammar of Spoken and Written English:-

a. **Non clausal units in conversational discourse**
Conversationaldiscourses make **wide use of non-clausal units**i.e.utterances which do not contain an explicit subject or verb.

b. **Personal Pronouns & Ellipsis in Conversation** Wide usage of **personal pronouns and ellipses in conversational discourse**. It is because of shared context in which conversation take place.

c. **Situational Ellipsis in Conversation** Speakers at times **intentionally skip certain words** as they take it for granted that the listener understands the immediate context.

d. **Non-clausal Units as Elliptic Replies in Conversation**
In the **shared social situation in which the conversation is taking place** both speakers know what is being talked about.

e. **Repetition in Conversation** **In order to give added emphasis** to a point the spoken discourse has abundant repetitions in conversation.

f. **Lexical Bundles in Conversational Discourse** There is frequent use of lexical bundles in conversational discourse such as

"**in addition to**", "**inorder to**","**it is going to be**","**If you want to**", etc.

8.   **Performance Phenomenon of Conversational Discourse**   Following are the important points of Performance Phenomenon of Conversational Discourse:-

a.   **Silent and Filled Pauses in Conversation**   Theindividuals fear to lose their turn if they pause in the end. **To retain their turn the speakers tend to pause** in the middle giving the impression of speaking the very next moment.

b.   **Utterance launchers and filled pauses** In   conversational discourse people use **utterance launchers such "as well as", "all right", "okay", "fine", "lets"**etc. to take the turn and to fill pauses.

c.   **Attention Signals In Conversation**   Speakers often **use other persons' names as an attention signal** to make it clear to whom they are speaking to.

d.   **Response Elicitors In Conversation**   **A question tag** is an example of **response elicitors** in conversation.

e.   **Non Clausal Items as Response Forms**   **Uh, Huh, Mum, Yeah, Okay, are the examples** of non-clausal items.

f.   **Extended Coordination Of Clauses**   In   conversational discourse one **clausal unit is added to another clausal unit** with such items "as,and, but etc."

9.   **Constructional principle of conversational discourse.**

There are key principles which underlie the production of conversational discourse, which are following.**The principle of keep talking, the principle of limited planning ahead and the principle of qualification** of what has been said.

a.   **Prefaces and Conversation**.   In conversation,the main part of speakers' message is oftenpreceded by a preface including fronting of **clausal units, noun phrase, discourse markers, linking adverbs, utterances launchers,interjection etc.**

b. **Tags in Conversation**. Speakers add tags in many ways to a grammatical unit in conversational discourse **by use of question tag at the end of a sentence to reinforce** what has just been said.

10. **Corpus Studies of the Social Nature of Discourse.** As **Swales** found spoken discourse to be **unpretentious in terms of vocabulary choice** while using the MICASE (Michigan Corpus of Academic Spoken English) he also found that spoken discourse **avoided name dropping** and the **use of obscure references**. He concluded that from the language point of view there are few barriers to cross - **disciplinary oral communication** than there perhaps might be in written academic communication.

11. **Collocation and Corpus Studies** Corpus studies are also used to examine collocation in spoken and written discourses for example **Hyland and Tse's 2004** study of dissertation acknowledgement of collocation "**special thanks", Sincere thanks" and "deep thanks".** They found that through this corpus how the writers expressed gratitude in their conversation.

12. **Criticism of Corpus Studies** Corpus study has been **severely criticized** by linguists saying that it is machine based orientation of Corpus Studies that leads to **atomized and misleadinginvestigation**of language use. Another criticism is that corpus studies do not take into account the **contextual aspects** of text.However, these arguments have been **negated by Tribble** saying that corpus studies enable contextual features such as the **social context of the text, communicative purpose of the text, roles of the readers and writers of the text shared culture** etc.