

IV. The Internal Structure of Discourse

a. Units of Analysis

Discourse analysts have always been looking for the unit of analysis that should be analyzed. Hence, a possible initial research question in such case might be simply formulated as: What are my units of analysis going to be? Which one of the following should I consider as a unit to be dealt with in discourse analysis?

Sentences

Utterances

Propositions

Turns

Speech acts?

Depending on the level on which they focus, researchers and discourse analysts may deal with larger or smaller units. Different approaches work with different units, but the same analyst may handle different units at the same time if s/he considers it appropriate for the purposes of his/her study (Alba Juez, 2009). According to Lakoff (1990), the term discourse is used to cover all linguistic interactions that follow predictable patterns known implicitly or explicitly to participants and which have a discernible function. In other words, Lakoff (1990) supports the view of discourse as linguistic interaction covering any length then the units chosen for analysis will also be greater or smaller depending, among other things, on the type of discourse used as data; for instance, a simple “Hi!” or “Okay!” (in a given context and situation) may constitute data for analysis.

In order to deal with discourse types one should talk about discourse dimensions. Discourses typically differ along a number of dimensions. These dimensions can be classified as the following

By means of production

By the medium of production

By the type of content (genre)

By style and register of the discourse

b. Discourse Dimensions

b.1. Means of Production: monologue and Dialogue

This dimension or classification is defined by the number of speakers involved in a particular discourse. Accordingly, two types of discourse in this case can be identified, a monologue and a dialogue. As far as a monologue is concerned, it is a discourse which is produced by a single speaker (writer), yet a dialogue is a discourse that is produced by more than one speaker (writer).

A dialogue can be a conversation or an exchange, as it can be called a turn taking which is composed of turns; this last can be considered as one of discourse units that can be considered in discourse analysis. One conversation (turn taking) may contain two turns and more.

Example: A: Are you going home?

B: Sure, I will be leaving in ten minutes.

A: Great!

In this dialogue A has two turns while B has only one. A widely used discourse analysis unit is the turn, which is commonly defined as one or more streams of speech bounded by speech of another, usually an interlocutor. According to Crystal (1969) and Coulthard (1985), the previous

definition of the unit ‘turn’ can be applied to the term ‘utterance’. However, some scholars such as Long (1980) and Crookes and Rulon (1985) claim that one turn may include a number of utterances since utterance is a stream of speech which is under one intonation boundary, bounded by pauses and constituting a single semantic unit. (more details will be discussed in the next chapter about Conversation Analysis).

b.2. Medium of Production: Spoken and Written

The distinction between speech and writing is often referred to as *channel* (Hymes,) or *medium* as speaking and writing involve different psychological processes. Spoken and written discourse differ for many reasons; in order to make a clear difference between both types of discourse a number of criteria should be followed such as grammatical intricacy, lexical density, nominalization, explicitness, contextualization, spontaneity, repetition and hesitation and redundancy.

- **Grammatical Intricacy:** Written discourse is more structurally complex and more elaborate than spoken discourse. That is, sentences in spoken discourse are short and simple, whereas they are longer and more complex in written discourse. Additionally, the use of passive voice is more frequent in written discourse than in spoken discourse. In written discourse we often use passive when we don’t want to specify the agent while in spoken discourse we would use a subject like “people”, “somebody”, “they”, “you” etc.
- **Lexical Density:** It refers to the ratio of *content words* (such as nouns, verbs, adjectives, and adverbs) to *grammatical or function words* (such as pronouns, prepositions, articles) within a clause. Spoken discourse is less lexically dense than written discourse. Content words tend to be spread out over a number of clauses, whereas they seem to be tightly

packed into individual clauses. In fact, spoken discourse has more pronouns, more lexical repetitions, more first person references, and more active verbs.

- **Nominalization:** It refers to presenting *actions* and events as *nouns* rather than as *verbs*. For instance, written discourse has a high level of nominalization: i.e. a large number of nouns is used compared to the use of verbs; yet, written discourse tends to have longer *noun groups* than spoken discourse.
- **Explicitness:** generally writing is more explicit than speech; however, this is not always the case since it depends on the purpose of the text. A writer/speaker can state something explicitly or infer it.
- **Contextualization:** it refers to the extent knowledge of context which is needed to interpret a text. Writing is more decontextualized than speech, in other words, speech is more attached to context than writing because speech depends on a shared situation and background for interpretation. However, some types of written discourse may show high dependence on shared contextual knowledge, for instance, personal letters between friends.
- **Spontaneity:** Spoken discourse lacks organization and is most of the time ungrammatical because it is spontaneous, whereas written discourse is organized and grammatically correct. Furthermore, spoken discourse contains uncompleted and reformulated sentences since speakers may interrupt and overlap.
- **Repetition, Hesitation and Redundancy:** Spoken discourse contains more repetition, hesitation and redundancy since it is produced on the spot (in the real time). Moreover, spoken discourse has many pauses and fillers such as *emmm, ehh, er, well, you know, I mean, etc.*