



A comparative study on the frequency of the usage of cataphora and anaphora in English and Persian narratives

Rana Azarizad and Iman Tohidian

Department of Foreign Languages, Iran University of Science and Technology (IUST), Tehran, Iran.

ARTICLE INFO

Article history:

Received: 22 January 2012;

Received in revised form:

17 February 2012;

Accepted: 26 February 2012;

Keywords

Anaphora,
Cataphora,
Endophoric,
Exophoric,
Narrative Essay.

ABSTRACT

Writing through an organized model and using the most frequent structural patterns as the proficient writers in different sorts of writings has always been worthy of attention. It can be said that narratives are nearly the most common and prevalent sort of writing in Iran. From the primary school, students are being asked to jot down their experiences of a specific occasion or holiday (e.g. Norouz holiday, summer holiday); however, little attention has been paid to the way students have to organize their writings and the kind of referencing has been taken for granted. The connectedness and cohesion of any text in all languages is granted, not mention the fact that coherence is the prerequisite to that. Using cohesive devices, which referencing is one of them, helps the writer to keep the cohesion. Looking for the differences in the commonality of the existence of cataphora and anaphora, we are going to come up with the most prevalent type of referencing, in order to help students to write in the most common way. Cataphora is usually defined as the referential relation in which the element referred to is anticipated by the referring element, usually a pronoun, and anaphora is a linguistic relation between two textual entities which is defined when a textual entity (the anaphor) refers to another entity of the text which usually occurs before (the antecedent). Two English and Persian contemporary short stories have been waded through to discover the most common pattern in terms of using anaphora and cataphora referencing in English writing. Ten English narrative essays and ten Persian narrative essays have been analyzed too for the same purpose. In the end, the results tend to show that anaphoric referencing comparing to cataphoric one is more prevalent in Persian narratives comparing to English ones. Of course, if we consider the frequency of anaphoric and cataphoric referencing, anaphora is more common in both languages; however, it is more common in Persian, as we see rare samples of cataphoric referencing in Persian, while there are more samples of cataphoric referencing in English narratives.

© 2012 Elixir All rights reserved.

Introduction

Reference in the systemic framework

In their influential book *Cohesion in English*, published in 1976, Halliday & Hasan state what they mean by their use of the term reference: "There are certain items in every language which have the property of reference, that is to say, instead of being interpreted semantically in their own right, they make reference to something else for their interpretation. In English these items are personals, demonstratives and comparatives" (p. 31). According to them, this class of items indicate the way the information is to be retrieved from elsewhere in the text. What is idiosyncratic in reference as a cohesive device is not the process of retrieval itself but, as the authors put it, "the specific nature of the information that is signalled for retrieval", that is "the identity of the particular thing or class of things that is being referred to" (p. 31).

After this general characterization, Halliday & Hasan establish a categorization of Reference according to the way it is activated in the English language. They establish a separation between two different cases of referencing (Exophora and Endophora) and first one refers to as situational, and the second as textual. As the authors put it (p. 33): "Exophora is not simply a synonym for referential meaning. Lexical items like John or

tree or run have referential meaning in that they are names for something: object, class of objects, process and the like". An exophoric item, does not name anything, and it signals that reference must be made to the context of situation. Both exophoric and endophoric reference embody an instruction to be interpreted from elsewhere that the required information is for interpreting the passage in question. What is essential to every instance of reference whether endophoric (textual) or exophoric (situational) is that there is a presupposition that must be considered. The main difference between Exophora and Endophora, for the purpose of the system of cohesion, is that Exophora is not cohesive.

By linking the language to the context of situation, it helps the creation of text, but it does not help the establishment of links between different passages in a text. It means that, 'exophoric reference does not contribute directly to the implementation of cohesion in a text, since it does not contribute, as the authors put it, to the INTEGRATION of one passage with another so that the two together form part of the SAME text" (p. 37). Endophora, on the contrary, is cohesive, as it contributes to the texture of a text, that is, it helps establishing links between different passages of a text. As a cohesive device, endophoric reference establishes itself in two different ways,

depending on whether its signalling is to previous or subsequent text. When the signalling is to previous text, reference is said to be anaphoric, and when it is to subsequent one, reference is said to be cataphoric. In *An Introduction to Functional Grammar*, Halliday (1986, 1994) does not introduce any change in this categorization of the system of cohesion. On the contrary, he further says that Reference first evolved as an exophoric relation and that, "more often than not, in all languages as we know them", items that may be used exophorically (e.g. the category of person and particularly the third person forms *he, she, it, they*) are used anaphorically, or, we would rather say, endophorically". (Halliday, 1994, P.95).

Furthermore, by stressing that, in the case of Reference, both cohesion and structure contribute to the quality of texture, Halliday definitely links endophoric reference also to structure, which being important for English is particularly vital for languages other than English, as we shall see: "The quality of texture depends partly on cohesion and partly on structure. If the pronoun and its referent are within the same clause complex, this is already one text by virtue of the structural relationship between the clauses; the cohesion merely adds a further dimension to the texture. If on the other hand there is no structural relationship, the cohesion becomes the sole linking feature, and hence critical to the creation of text. The cohesive relationship itself is not affected by considerations of structure; Peter... he form an identical pattern whether they are within the same clause complex or not. But they carry a greater load in the discourse if they are not" (Halliday, 1994, P. 312).

The linguistic phenomenon of anaphora has been under the attention of linguists and computational linguists for some time now, leading to different interpretations. Basically, anaphora is the phenomenon of reiteration of an entity (called 'antecedent') by a reference (called 'anaphor') that points back to that entity. For practical reasons, we will call 'referential expressions' (REs) both participants in an anaphoric relation. Actually, during the reading of a text, it is very likely that an anaphor become, in its turn, an antecedent for another co-referential anaphor that follows it. In such case we will use the term 'anaphoric chain' to denote the textual relation of the co-referential REs.

The process of identifying the antecedent of an anaphor is called 'anaphora resolution' (AR). Its automatization represents one of the main jobs of computational linguists, as anaphora resolution is so important. In order to develop an anaphora resolution system, it is necessary to fully understand the nature of the referential process in discourse and the problems behind it. Since 1976, Halliday and Hassan stressed that the anaphoric relation is a semantic, and not a textual one. Although it is somehow agreed that semantic features are essential for anaphora resolution, the authors of automatic systems devised so far preferred to avoid the broad use of semantic information (Lappin & Leass, 1994), (Mitkov, 1997), (Kameyama, 1997). This choice, motivated by the difficulty and complexity of achieving a correct semantic approach, has had as this result that an algorithm of anaphora frequency with a very high degree of success has not been found yet.

Discourse is more than mere words

Words are much more than only conventions that pick things from the real world. Starting with the school of Ferdinand de Saussure, linguists have acknowledged the conceptual nature of words. Saussure's famous dichotomy defining the linguistic sign (signified/significant) introduces the idea of a concept present behind every word in the mind of a speaker. Both

Saussure and the linguists that have followed and developed his theories do not consider the entities from the real world as playing any part in the act of speaking. In this regard, Hjelmslev (1961) stresses that the nature of the signified is a purely mental one. A different approach, which of Ogden and Richards (1923), leads to a similar conclusion. Though the semiotic triangle proposed does include, as a third component, a material referent, its role is considerably minimised. Moreover, in analysing the three components of the triangle (symbol, reference, referent), (Ullmann, 1962) states that the study of the referent and its relation with the reference should be the concern of philosophy or other sciences, while linguistics should only deal with the reference-symbol relation.

"To complete the set of semantic characteristics attributed to a discourse entity (DE) it is imperious to have in view the co-referentiality of notions in a discourse. In this respect, endophoric references, in the form of anaphora/cataphora, are essential. Halliday and Hassan consider the anaphoric function as crucial in creating cohesive links within sentences. As we share their opinion that the reference items must match the semantic properties of the item referred to we believe that an anaphora model should necessarily take into account the semantic representation of the words involved in discourse" (Halliday and Hassan, 1976, P. 32).

Cataphora

Cataphora is usually defined as the referential relation in which the element referred to is anticipated by the referring element, usually a pronoun. Scepticism about the existence of cataphora has been expressed by a number of scholars such as Bolinger (1977), Stockwell (1995), Cornish (1996), etc. Sometimes this scepticism has even embraced a categorical denial. For instance Kuno (1972), who believes that all seemingly cataphoric pronouns must have their co-referential expressions somewhere in the preceding text.

This firm position is refuted by Carden (1982), as Tanaka (1999) notes, which provides evidence, within the 800 examples he collected from corpora, of pronouns that represent the first mention of the discourse entity in the text. Examples of the type "when he realises something, X (= he) does something else" are frequent in newspaper articles and television programmes:

When he became president, George Bush renewed his appeal for a "kinder, gentler nation." (Compton's *Interactive Encyclopedia*, 1995, title word *Bush*)

Gordon and Hendricks (1997), found that pronoun – name co-reference is frequent in situation when subordinate clause or a prepositional phrase, containing the pronoun, precedes the main clause, containing the co-referent noun. Also in literary texts writers often prefer to introduce a character with a pronoun, which has a sum of characteristics, attached, and only later his name is mentioned: "From the corner of the divan of Persian saddle/bags on which he was lying, smoking, as was his custom, innumerable cigarettes, Lord Henry Wotton could just catch the gleam of the honey-sweet and honey-coloured blossoms of a laburnum..." (O. Wilde – *The Picture of Dorian Gray*)

The connectedness of any text in all languages is granted, in the first place by:

- a) Lexical connections, i.e. either repeating the same lexical items or using lexical items of the same or related semantic groups to make the text semantically homogeneous.
- b) Anaphoric cohesive devices: anaphora and cataphora.
- c) Substitution cohesive device.

The semantic basis for the general anaphoric cohesive device universally employed in text building is formed with co-reference to the identical meaningful units by nominal headwords (antecedents) on the one hand, and pro-forms (anaphoric elements) on the other, which creates close ties between the items of text, and the whole pattern of anaphoric device is mostly an utterance dependent endophoric structure. Particular kinds of the general anaphoric device are anaphora and cataphora. Both of them are patterns of the matization. The structuring of anaphora and ordering of its items follows the pattern of a natural logical (and linguistic) sequence of text items in discourse – antecedent nominal first, and a pro-form following it.

Because of such natural order anaphora is the most widely used cohesion pattern – a neutral one, and with anaphoric element devoid of its own content with the only function of connecting parts in text; that is why anaphora is hardly ever perceived by either speaker or listener. E.g. *Our teacher is very kind; he never criticizes us for little spelling errors.*

Cataphora is a variant of the anaphoric device which also rests on identical co-reference of cataphoric pronominal element and a nominal headword, but ordering in cataphora is different from anaphora: an inversion of natural logical sequence with a pro-form occurring first and a nominal headword following is the pattern order of cataphora, like in:

The one to start discussion was my son!

Go there, and see for yourself. Everybody is in the kitchen

Such contrary-to-logic ordering of co-referring units (it could also be called front-referring) naturally causes semantic markedness of this pattern with certain additional specific meanings such as, e.g., emphasis or modality.

Anaphora

Anaphora is a linguistic relation between two textual entities which is defined when a textual entity (the anaphor) refers to another entity of the text which usually occurs before (the antecedent). When the anaphor refers to an antecedent and both have the same referent in the real world, they are called coreferential. Although, coreference and anaphora are two different concepts, in reality, they most often co-occur except in some cases. Note that, not all varieties of anaphora are based on referring expressions such as verb anaphora in the first example or bound anaphors in second example. On the contrary, coreference may occur without anaphora. For example, the use of the same proper name consecutively with each one referring to the same entity. The anaphora is defined as being the resumption of an entity exists previously in the text, whereas, the cataphora occurs when a reference is made on an entity mentioned further in the text.

Anaphoric expressions linguistic categories

Pronominal Anaphora: Pronouns form a special class of anaphors because of their empty semantic structure. They do not have an independent meaning from their antecedent. In addition, in Persian, not all pronouns are anaphoric: e.g., deictic pronouns such as “*من*” ‘I’, “*تو*” ‘you’ and “*ما*” ‘we’ are not anaphoric ones. Pronominal anaphora includes third personal pronouns (غایب ضمیر), demonstrative pronouns (اسم های اشاره) and relative pronouns (اسمهای موصولی)

Personal Pronouns: In Persian, third personal pronouns can be classified in disjoint or joint pronouns and also in nominative, dative or accusative ones. Thus, we distinguish:

- Nominative disjoint personal pronoun, like: او شما ایشان

Relative Pronouns: The relative pronoun in Persian is always anaphoric and is referring to the immediate previously mentioned noun phrase, like:

کسی که و کسانی که

Lexical Anaphora: Lexical anaphora is realised when the referring expressions are definite descriptions or proper names. These definite expressions increase the cohesiveness of the text and moreover they convey some additional information (synonymy, generalization, specialization).

Verb Anaphora: Verb anaphora is another variety of anaphora which is characterized by the use of the verb (do فعل).

(ما زنده ایم تا به تعهدات خویش عمل کنیم)

“We live to do our obligations”

Compared with other languages such as English or French, anaphoric expressions in Persian are almost classified similarly, although, some particularities to the Persian language can be noticed. On the one hand, the third person pronouns can be used as demonstrative pronouns (e.g., او رئیس جمهور است) He is the president.

“In some cases, the anaphora resolution moment may be delayed until other discourse elements intervene to elucidate the anaphoric co-reference. This is the case in the following example” (Tanaka, 1999, P. 221):

Police officer David Cheshire went to Dillard's home. Putting his ear next to Dillard's head, Cheshire heard the music also.

The disambiguation moment of the pronoun *his* is the moment the reader processes *Dillard's head*. An inference allows the recuperation of [Cheshire] instead of [Dillard], since they were the only characters in the story and, by pragmatic knowledge, the system should recognize that a man cannot put part of his head next to the head itself. Therefore, the resolution moment is not that of the pronoun reading, and neither that of the succeeding co-referential proper noun reading, but intermediate. The proper noun reading will, perhaps, strengthen the belief that the antecedent is [Cheshire], as inferred.

Also, “sometimes the pronoun resolution may be based on an inference without an explicit restating” (Tanaka, 1999, P. 252):

The government contended Jacobson, 48, former big-time horse trainer turned East Side real estate operator, killed Tupper because Miss Cain, his live-in girlfriend of five years, moved from his apartment to Tupper's just down the hall.

It becomes obvious from these examples that pronoun anaphora resolution might not be immediate even in the case of classical anaphora. When a reader cannot be sure about the antecedent at the encounter point (*his*), the final resolution will be deferred until he has enough information to disambiguate it from the later input (in this case, *Tupper's*). It is therefore correct to assume that, in general, there is a distinction between the point when a reader encounters a pronoun and initiates its interpretation (initiation point), and the point when the reader completes the interpretation of the pronoun (completion point). As Sanford and (Garrod, 1989) note, the gap between the two points can be nil, when a reader resolves a pronoun at the very moment he encounters it, or it can be extended to the end of the phrase, clause, or sentence in which the pronoun is included. During the gap between the initiation point and the completion point, humans retain the information obtained by processing the co-text of the pronoun in some kind of temporary location of memory until the resolution is stabilised.

Narrative Writing

Narration is a rhetorical mode which tells the story of what has happened in the past in a chronological order. Here the writer arranges the actions based on the precedence in time. It doesn't show any cause and effect or any classification and the like. Due to the sequencing of events based on the time they have happened, this mode of writing is probably one of the easiest types and generally students are better at writing this type of essay in comparison with other types.

According to Hutchinson (2005) "when you write a narrative, you are telling a story.

Who's a better audience for your story than a friend?

One way we tell our friends stories are by writing letters. Good topics for letters include the interesting things that happen to us during the day."(P.73)

Research has shown that the most familiar and the simplest mode of writing is the narration. The definition of narrative text is not an easy one accepted by all researchers and many definitions have been provided by the difference to scholars based on their specific views toward a narration structure. Narratives may report the actual or fictitious about real people and real worlds or about imaginary people and imaginary worlds. They include categories like myth, epics, fable, folk tales, short stories novels tragedy and comedy.

In linguistic studies, a narrative is considered to be a text in which the speaker or writer relates a series of real or fictitious events in the time order in which they took place. Richards and Schmidt (2002) define this writing mode as 'narrative writing reports an event or tell the story of something that happened'. The purpose of a narrative text is to inform, entertain and excite the readers. Narrative text depicts events, actions, emotions, or situations the people in a culture experience.

The chronological order of events which are organized around a central conflict is the foundation of a narrative text. According to each narrative essay has a central conflict which is developed during the time. In fact, such a conflict which can be between man and nature, between man and man, or between man and himself presents a series of crises and the climax of the narration reveals that which side of the conflict has won.

Narration is different from storytelling because narration has a broader meaning. Narration is the telling or retelling of what has happened over duration of time and the causes and effect relations are not involved.

Method

Procedure

Two contemporary English and Persian short stories and ten narrative essays in English and ten in Persian were waded through to discover the most prevalent pattern in terms of using anaphoric and cataphoric referencing in English writing.

Data source

The English short stories named Mr. 'Washington', by Les Brown, and 'Faith, hope and love', by Peter Spelke were analyzed. Also two Persian short stories were selected from Hossein Shekarriz, named 'A gift full of love' and 'Generosity'. Ten English narrative essays were selected from the book *Composing with confidence* (Mayers, 2006). The Persian essays were selected from the book *'the stories of 1001 nights'* by Mousa Farhang.

Instrument

The main instrument used for the fulfilment of this research is the number of occurrence of an element in the context, namely the frequency of that item.

Results

After analyzing the two English and Persian short stories and the ten English and Persian narrative essays, we came up with Figure 1. As it is shown in the following table, the use of anaphoric referencing is more common in both English and Persian narratives and short stories; however, the cataphoric referencing is not that frequent in Persian narratives, comparing to that of those in English narratives.

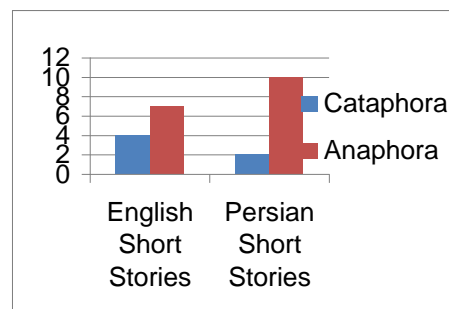


Figure 1. Statistical study of endophoric referencing in two short stories

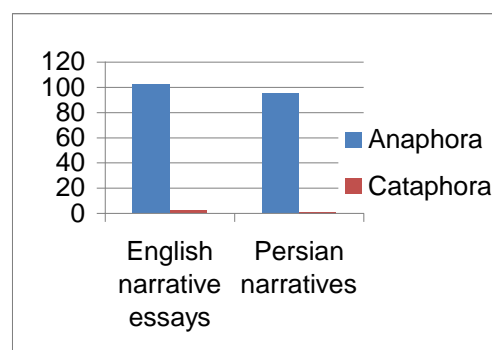


Figure 2. Statistical study of endophoric referencing in ten narratives

An interesting case of cataphoric referencing in both English and Persian was 'numbering the reasons'. E.g., 'there are some reasons that I don't like your nation:'. As you see in this example, first the word reason has been mentioned and then the reasons are named (Crystal, 2003 and Richard and Schmidt, 2002). This is again an example of cataphoric referencing.

Conclusion and Discussion

Through this paper there was an attempt to verify the following hypothesis: 'Using cataphoric referencing seems to be more frequent in Persian narratives and short stories, comparing to that of those in English.' After analysing some English and Persian narrative and short stories, the researchers came up with this conclusion that the use of anaphoric referencing is more common in both English and Persian narratives and short stories; however, the cataphoric referencing is not that frequent in Persian narratives, comparing to that of those in English narratives. Also Anaphoric referencing is more frequent in English narratives than Persian narratives. In this case it can be said that the purpose of this paper which was investigating the frequency of using the anaphoric and cataphoric referencing in two types of writing in English, namely narrative essays and short stories has been fulfilled through the statistical analysis. On the one hand, the number of the texts which have been analyzed has been limited to ten English and ten Persian narratives, also two Persian and two English short stories. But on the other hand there was not an exact equivalent of English narratives essays in Persian, the researchers have delimited that

to an old Persian story (10001 nights) that is the closest equivalent to English narratives.

Most of the students are taught to learn the format of an essay and well-formed grammatical structures to be used in their writing; however, little attempt has been done to make students familiar with the most prevalent pattern in referencing. Through this statistical research, the attempt has been done to reach the most frequent pattern in using anaphoric and cataphoric referencing, especially in narrative writing which is the easiest type of writing especially for lower level students.

References

- Bolinger, D. (1977). *Pronouns and Repeated Nouns*. Indiana: Indiana University Linguistics Club.
- Carden, G. (1982). Backwards anaphora in discourse context. *Journal of Linguistics*, 18, 102-137.
- Cornish, F. (1996). Antecedentless anaphors: Deixis, anaphora, or what? Some evidence from English and French. *Journal of Linguistics*, 32, 83-102.
- Crystal, D. (2003). *A dictionary of linguistics and phonetics*. Melbourn: Blackwell Publishing.
- Gordon, P.C., & Hendrick, R. (1997). Intuitive knowledge of linguistic coreference. *Cognition*, 62, 27-48.
- Halliday, M.A.K., & Hassan, R. (1976). *Cohesion in English*. Longman: London.
- Hjelmslev, L. (1961). *Prolegomena to a Theory of Language*. Madison: Wisconsin University Press.
- Hutchinson, E. (2005). *Narrative writing*. USA: SADDLEBACK.
- Kameyama, M. (1997). Recognizing Referential Links: an Information Extraction Perspective, *Proceedings of a Workshop*

Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts, Madrid.

- Kuno, S. (1972). Functional Sentence Perspective: A Case Study from Japanese and English. *Linguistic Inquiry*, 3, 36-68.
- Lappin, Y., & Leass, H. J. (1994). An Algorithm for Pronominal Anaphora Resolution in *Computational Linguistics*, 20, 69-106.
- Meyers, A. (2006). *Composing with confidence*. New York: Longman.
- Mitkov, R. (1997). Factors in Anaphora Resolution: They are not the only things that matter. In Mitko R., & Boguraev B. (Ed.) *Proceedings of the Workshop Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts*. Madrid: Universidad Nacional de Educación a Distancia.
- Ogden, C.K., & Richards, I.A. (1923). *The Meaning of Meaning*. Londra: Routledge and Kegan Paul.
- Richards, J.C., & Schmidt, R. (2002). *Dictionary of Language teaching and applied linguistics*. Longman: London.
- Sanford, A.J., & Garrod, S.C. (1989). What, when, and how?: Questions of immediacy in anaphoric reference resolution. *Language and cognitive process*, 4,(3), 82-101.
- Stockwell, P. (1995). How to Create Universes with Words: Referentiality and Science Fictionality, *Journal of Literary Semantics*, 23, 14-56.
- Tanaka, I. (1999). *The value of an annotated corpus in the investigation of anaphoric pronouns, with particular reference to backwards anaphora in English*, PhD. thesis, University of Lancaster.
- Ullmann, St. (1962). *Semantics: an Introduction to the Science of Meaning*. Oxford: Blackwell.