Sentential Effects on Discourse Processing

Mohamed Taha Mohamed (*)

Abstract

While there is an established literature on the effect of discourse and context on sentence processing, little is known about the effect of sentence on the text. The current study reviews recent developments in sentential variables' effect on discourse processing. Four sentential variables were discussed: anaphoric referentiality, implicit causality, syntactic cues, and linguistic connectives and markers The study aim at presenting an integrative framework of how these variables set up and update the reader's mental model of the text. Sentential variables build this mental model through connecting consequent sentences, avoid repeating the same referent, specifying the focus of the text, and keeping track of the temporal order of events and logical relations among them. In discussing each of these variables, basic aspects of the variable was presented and new developments in its linguistic analysis and related psycholinguistic evidence were reviewed. Controversial issues like time course of the effect of these variables, and problems of the relative weights assigned to them were mentioned also in order to locate points of possible interest for future research.

Keywords: anaphoric referentiality- implicit causality- linguistic markers

*Department of Psychology - Ain Shams University

تأثير متغيرات الجملة على معالجة النص

محمد طه محمد

ملخص

بينما يوجد تراث مستقر حول تأثير النص والسياق على معالجة الجملة، فان هناك القليل مما هو معروف عن تأثير الجملة على معالجة النص. وتراجع الدراسة الحالية التطورات الحديثة فى تأثير متغيرات الجملة على معالجة الخطاب. وقد تمت مناقشة أربعة متغيرات فى هذه الدراسة: مرجعية الضمائر، و السببية الضمنية، والمهديات التركيبية، والروابط و المؤشرات اللغوية. و تهدف الدراسة الى تقديم اطار متكامل حول كيفية قيام هذه المتغيرات بتأسيس و تحديث النموذج العقلى للنص لدى القارىء، وذلك من خلال ربط الجمل المتعاقبة، وتحديد بؤرة النص، و متابعة التتابع الزمنى للأحداث والروابط المنطقية بينها. وعند مناقشة كل متغير من هذه المتغيرات، فقد تم عرض الجوانب الأساسية للمتغير بالإضافة إلى مراجعة التطورات الحديثة فى تحليله اللغوى والدراسات النفسية اللغوية ذات الصلة به. و تمت الإشارة الى بعض القضايا الخلافية مثل المسار الزمنى لتأثير بعض المتغيرات والوزن النسبى لها لكل منها، وذلك بغرض تحديد بعض النقاط التى قد تكون مثار اهتمام للبحوث المستقبلية.

1. Overview

There is strong evidence from psycholinguistic research, supported by common sense, that a sentence can't be fully understood without considering the text and/or the context in which it appears. That is, the whole text or situation affects the way we understand a single sentence. Different contexts can lead to different interpretations of the same sentence. This claim is shared by researchers from disciplines as different from each other as computational linguistics and literary criticism, including sociolinguistics and pragmatics, psychology of thinking (the notions of scheme and script), social cognition, and decision-making. However, a little is known about the effects a sentence, or sentential variables, might have on processing a discourse .

The current study is an attempt to present the major, recent trends of sentential interaction with discourse, basically as conceptualized in psycholinguistics. More specifically, the study seeks to present an integrative framework of how sentential linguistic characteristics affect the processing of an extended text. In other words, it is an attempt to define sentential variables that affect the construction of discourse representation, as proposed in psycholinguistic literature. These variables include anaphoric referentiality, implicit verb causality, syntactic cues, and linguistic connectives and markers. They are critical for both setting up a mental model of the discourse and updating this model as the reader continues reading the text. In discussing the contributions of these variables to discourse comprehension, recent developments in linguistic analysis of these variables and empirical evidence about their roles in text comprehension will be presented.

Studying these variables was motivated by recent developments in theories of discourse comprehension. That is, these theories from almost all frameworks give special attention to tracking how a new sentence can change the discourse representation that was set up until the previous sentence. This can be seen in cyclical processing of the text in both bottom up approach (e.g., Kintsch, 1998; Kintsch & Van Dijk, 1978) and in linguistically-oriented approach (e.g., Heim's (1982) file change theory). Similarly, the output could be a new

matrix of connection strengths with each new sentence as in computational modeling (e.g., Chater & Christiansen, 1999), or a new layout of discourse management as in Fauconnier's theory of mental spaces (e.g., Fauconnier & Turner, 1998). (For a revision of discourse comprehension theories, see Mohamed, 2007).

However, in characterizing these sentential effects on discourse representation, some psychological models (like bottom-up ones) give more weight to low level connections among parts of the text (through spread of activation mechanism) and don't give much consideration to linguistic determinants of these connections, such as the previously mentioned sentential variables. Even for these approaches that don't give much weight to low level associationistic connections, like mental model framework, they fail to specify how their models are built out of these sentences or how sentential variables can maintain or change the model. A basic assumption adopted here is that specifying sentential, linguistic variables and integrating them within a low level, connectionistic mechanisms might lead to a substantial progress in comprehending the mechanisms of discourse processing.

Thus, the current study is an attempt to review and integrate these linguistic variables and to discuss how they might affect discourse comprehension. To achieve this goal, the contribution of each of the four linguistic sentential variables mentioned above (anaphoric referentiality, implicit causality, syntactic cues, and linguistic markers) in text comprehension will be discussed in the next four sections, respectively, in terms of linguistic analysis and empirical evidence. A concluding section comes at the end of the study to indicate the major findings and their significance .

2. Anaphoric Referentiality

Anaphoric referentiality is a linguistic mechanism that maintains the coherence and continuity of discourse. Among other things, it helps speaker/writer to avoid overt repetition of the same referent over and over. On the other hand, it helps the reader/listener to arrange information about someone or some event as a unit instead of dealing with a set of separate, fragmented sentences. According to bottom-up models of discourse comprehension (e.g., Kintsch & Van Dijk, 1978), anaphoric referentiality to referents in the sentence depends on

automatic, low level process of argument overlap. However, mental model theorists (e.g., Garnham, 1999; Johnson-Laird, 1983), believe that anaphor resolution can't be understood outside the representation of the situation described in the discourse. In other words, according to this framework, anaphor resolution is a search for a referent in a mental model and the result of this search is determined by context and world knowledge, rather than by serial search among possible antecedents.

This section focuses on the effect of anaphoric referentiality on discourse comprehension. The process of anaphor resolution or the factors that might affect it will be mentioned in relation to this goal. To organize the presentation, anaphoric effects can be divided into two dimensions: forms of anaphora and levels of anaphora. Both dimensions will be discussed below.

Forms of anaphora

It is reasonable to expect that different forms of anaphora refer to different types of discourse referents and consequently have substantial effect on discourse comprehension. A major distinction in this regard is between definite and indefinite anaphoric reference. For a long time, it has been assumed that definite description refers to a referent that was already mentioned before while indefinite description introduces a new referent for the first time in the discourse (as in Heim's (1982) theory).

The significance of this distinction for discourse comprehension can be thought of in the light of an earlier distinction between given and new information that was proposed by Haviland and Clark (1974). According to Haviland and Clark, what makes a discourse coherent is the easy integration between the given and new information in it. That is, with each new sentence, the reader distinguishes between given, presupposed information, on one hand, and new, inserted information, on the other hand. Then, s/he tries to match or relate given information with the information he has in memory before inserting the new information. Only then, s/he can relate new information to old one in a new, integrated knowledge structure. If the reader fails to find this old information that matches the given information in the sentence, he will have to build that matching by generating a bridging inference.

This failure is time- consuming. Accordingly, the target sentence "the beer was warm" will be read faster after "we got some beer out of the trunk" than after "we checked the picnic supplies" (examples are from Haviland & Clark, 1974). That is because given information in the target sentence "the beer" is easier to be matched with the context of the former sentence than the latter one. The matching between given and new information as a basis for discourse integration might clarify the relation between definite and indefinite anaphors. Definite anaphora describes the "given" information whereas indefinite anaphora describes "new" information. So in sentences 1 and 2 below, the new description in the sentence 2 should be linked to previous, given information in sentence 1 to indicate the identity of the two referents and, accordingly, to maintain the coherence of the discourse.

1. Mary saw a man$_i$ in the street .
2. He$_i$ was drunk.

Also, by referring to a previously mentioned referent, definite description saves the processing time and effort needed to build a new structure in the discourse or to decide whether to build a new one or not. Rather, the processor focuses on the new, inserted information and how it might be integrated in the mental model. Haviland and Clark considered this matching process necessary for acquiring new information .

The effect of anaphora form on discourse comprehension is explained in further details within two theoretical conceptualizations of reference and language comprehension: Locality of Syntactic Dependencies theory (Gibson, 1998) and Centering Theory (Gordon et al, 1993). The basic framework of each of theory will be briefly presented followed by its account of the effect of different anaphoric forms on discourse comprehension.

Gibson syntactic dependencies theory is a computational, activation-based theory of sentence processing that aims to specify the interaction of different constraints. It assumes that there is a limited pool of computational resources available at any time. Different language processes require using different quantities of these resources. To get to these computational resources, different structures try to reach a target activation threshold. If a structure was favored by

most constraints, it gets high level of activation, and accordingly, the time needed to comprehend it would be short. On the other hand, if different constraints favor different structures, there will be a competition among these structures to get the largest possible amount of computational resources. This competition will be reflected in increasing the time necessary to reach an interpretation of the sentence. In terms of representation, Gibson's theory assumes a ranked parallel processor. It keeps the representations with the highest activation (above a target threshold) in the focus. Other representations that have lower activation level than the target threshold, but equal or higher than a second threshold, will be kept in the active representation set. Other representations that have lower activation than the second threshold will be retained in inactive representation set. The theory has two components, integration cost component and memory cost component. In integration component, there is a fixed amount of computational resources necessary as a cost for any integration process but this cost increases in proportion to the distance between the integrated elements. Accordingly, local integration is easier than distant one. In memory component, it is assumed that there is a fixed amount of computational resources necessary to remember each syntactic category needed to complete structural analysis of the current input string. Assuming lexically-based predictions of syntactic categories, sentences with more intervening syntactic structures should require more computational resources, and take more time to be processed than sentences with less intervening structures .

According to the theory, computational resources are needed to build a structure for a new discourse referent, as indicated by indefinite noun phrase (NP). That is, processing these new referents causes more integration cost that eventually leads to a substantial decay in the activation assigned to preceding lexical items. To prove this hypothesis, Warren and Gibson (2002) had their subjects give acceptibility rates of sentences with doubly nested relative clause (RC)1 structures. The most embedded subject position in these sentences is occupied by either an indexical pronoun, proper name, full NP, or a pronoun with no referent, as in (3).

(3a) The student who the professor who I collaborated with had advised copied the article. (indexical pronoun, I)
(3b) The student who the professor who Jen collaborated with had advised copied the article. (proper name, Jen)
(3c) The student who the professor who the scientist collaborated with had advised copied the article. (full NP, the scientist)
(3d) The student who the professor who they collaborated with had advised copied the article. (no referent pronoun, they)

They found that (3a) was rated as significantly easier than the other three versions of the sentence. The results could be accounted for within Gibson's theory as follows: New NP referents as in (3b) increases the memory cost as they require building a new discourse structure (in addition to the syntactic categories predicted at this point). With definite old referent (3c), on the other hand, the referent is in the discourse structure and there is no need for building a new one. Of course (3d) is the most difficult because using no referent pronoun would require the reader to build bridging inferences to integrate this no referent pronoun with previous antecedent in memory.

If integration cost, found in the previous study, was related to building discourse structures for new referents, then integration across new referents should be more difficult than integration across non-referents. To test this hypothesis, Warren and Gibson (2000) compared subjects' acceptability ratings of doubly nested relative clause (RC) structures with the quantifier "no one" (where there is no referent) either in outer, middle, or inner position, as in (4).

(4a) *No one who the administrator who the teacher talked to commended was fired .*
(4b) The teacher who no one who the administrator talked to commended was fired.
(4c) The administrator who the teacher who no one talked to commended was fired .

The results showed that the sentences with outer and middle quantifiers (4a and 4b), where most of integration occurs across new referents, were rated to be more complex than the sentence with inner quantifiers (4c).

On the other hand, centering theory (Gordon et al, 1993) assumes that all semantic entities referred to in an utterance are discourse centers, which are linked together to make a discourse coherent. There is a distinction between two types of centers in an un-initial sentence. The first is backward-looking center (cb) that provides a link to the preceding utterance. The second type of centers is forward-looking centers (cf) that provide potential links to the subsequent utterance. Within this framework, (cb) of any utterance should be realized as a pronoun. That is because the pronoun here works as cue for connecting the two sentences. More specifically, it indicates that information in current sentence should be interpreted in relation to the information presented earlier in the previous utterance. Using a repeated name, according to the theory, deprives the reader from an important integration cue. Accordingly, in 5 below (taken from Gordon et al, 1993, p.313), sentence 5c (with cb realized as pronoun) is assumed to be better integrated with 5 b than sentence 5c` (with cb realized as repeated name).

(5a) *Susan gave Betsy a pet hamster.*
(5b) She reminded her such hamsters were quite shy.
(5c) Susan asked her whether she liked the gift .
(5c) She asked Betsy whether she liked the gift.

Empirical data supported the hypothesis and sentences with repeated names took more time to be read than sentences with pronouns. They called this effect the repeated name penalty. Subsequent research (e.g., Gordon & Chan, 1995) replicated this effect and showed that it occurs only for the grammatical subject of the sentence with agent rather than patient thematic role. Gordon and Chan (1995) compare repeated name penalty for active and passive versions of the same sentence as in (6).

(6a) *Susan decided to give Fred a hamster .*
(6b) She/Susan told him exactly what to feed it. (active)
(6b) She/Susan was questioned by him about what to feed it. (passive)

They found repeated name penalty for active and passive sentences though the grammatical subject played different thematic roles in the two constructions (agent in the active sentence but patient in the

passive one). However, repeated name penalty was more if the grammatical subject was agent than if it was patient. This was interpreted to mean that grammatical subject position is preferred as (cb) of an utterance. Again, the results supported the notion that the form of the reference (full NP vs. pronoun) affects discourse coherence and whether it would be easy or difficult to understand.

Despite this wide range of evidence, there might be other factors that need to be considered in accounting for the effect of anaphoric form on discourse comprehension. For example, Johnson-Laird (1983) questioned the generalization that definite anaphors refer to an already mentioned reference while indefinite anaphors introduce a new one. He emphasized the role of mental models and background knowledge in understanding the referents. So, he mentioned cases where definite description does not necessarily refer to an already-mentioned referent as in a sentence like "Ann was in a shop. She was talking to the assistant", where the definite description "the assistant" needs not to be previously introduced. Rather it can be inferred depending on the background knowledge of the prototypical shop. On the other hand, according to Johnson-Laird, indefinite description does not necessarily introduce a new referent. He gave the following example of short text (Johnson-laird, 1983, p.383), in which indefinite description can be introducing a single referent for the whole text or three separate referents depending on the previous knowledge (that it is the same person) .

> *"This is a story about a man who was war correspondent in the Boer war, a man who became prime minister, and a man whose wife burnt his portrait by Graham Sutherlad* ".

On the other hand, recent research by Almor (1999) indicates that the distinction between the pronoun and repeated names in centering theory might not be a clear-cut one. Almor's claim is that anaphoric phrases vary according to what he calls "information load". According to this hypothesis, anaphoric phrases reflect different amounts of information about their antecedents. A high information load anaphor (like repeated name) might be required for successful coherence if the antecedent is unavailable. However, this high information load might hurt the discourse coherence if the antecedent is relatively available

(as in focus of the preceding sentence), which might result in repeated name penalty .

Levels of anaphora

This section describes different effects of different levels of anaphora on discourse representation. Three sets of anaphora, with two levels in each of them, will be discussed: deep vs. surface, referential vs. attributive, and indirect vs. direct anaphora .

The distinction between *surface and deep anaphora* was first proposed by Hankamer and Sag (1976), who relabeled them later as ellipsis2 and model-interpretive anaphora (Sag & Hankamer, 1984). Two major differences can be identified between deep and surface anaphora. First, deep anaphora accesses the mental model level where they don't need a linguistic antecedent. Rather, antecedent is determined by non-linguistic context. Accordingly, if a person asks "who's she?" in a party, the other person will understand the antecedent of "she" even if it is not explicitly mentioned. This type of anaphora is assumed to be pragmatically controlled. On the other hand, surface anaphora requires explicitly mentioned linguistic antecedent and, according to the hypothesis, accesses pure linguistic model. Secondly, surface anaphora requires their antecedents to be presented in a form parallel to their own, while deep anaphora are not sensitive to the form of the antecedent. Accordingly, a deep anaphora like "do it" is not sensitive to the voice of the sentence as in (7) (from Hankamer& Sag, 1976), where (7a) and (7b) are both acceptable sentences.

(7a*) Someone had to take the oats down to the bin, so Sandy did it.*

(7b) The oats had to be taken down to the bin. So Sandy did it .

However, with ellipsis, the sentence will be acceptable only in active voice but not in passive voice, as in 8 and 9, respectively:

(8) *Someone had to take the oats down to the bin, so Sandy did.*

(9) The oats had to be taken down to the bin. So sandy did.

In an earlier test of this hypothesis, Murphy (1985) compared the effect of (1) length of antecedents, distance between the anaphor and its antecedent, and (2) the syntactic parallelism between the anaphora and the antecedent, on processing a sentence that contains deep and surface anaphora. According to Hankamer and Sage's hypothesis, the

effects of these two variables should differ depending on the type of anaphora. That is, while the length of antecedent effect should be the same for both types of anaphora, syntactic parallelism is necessary for surface anaphora but not for deep anaphora. The results showed that though these factors have significant main effects, neither of them interacted with the type of anaphora. That is, the effect of the controlled variables did not differ depending on the type of anaphora, as hypothesized by Hankamer and Sag. However, Tannenhaus and Carlson (1990) found differences of the effects of similar factors on deep and surface anaphora using acceptibility judgments. Similar results were obtained by others (e.g., Mauner et al (1995)) .

Another area where anaphors affect discourse differently is in the distinction between *referential and attributive anaphora*. In attributive anaphora, rather than referring to a particular individual or event as in referential anaphora, it refers to whoever satisfies the description (Johnson-laird, 1983). Distinguishing between referential and attributive anaphors may be determined by the context and the mental model. A description like "the winning captain" (Johnson-laird, 1983, p.385) is referential after the match, where it can be used to refer to a specific person. On the other hand, this description is attributive before the match. In this case, it is used to describe the person who will satisfy the description. Moreover, the same anaphor can be referential for somebody and attributive to the other depending on the mental model. Accordingly, in 10 (from Johnson laird, 1983), the definite anaphor is referential for the speaker but attributive to the listener, while it is the opposite in (11).

(10) *I don't want to tell you anything about the person I met yesterday .*

(11) *Double the number you're thinking of .*

This led Johnson-laird (1983) to assume that both speaker and listener has his/her model and that potential discrepancies enhance more work on the side of the speaker to make the description closer to the referent. Listener on the other hand, will use any available knowledge to construct his/her model.

The last type of anaphora to be discussed is the *indirect anaphora*. Two types of indirect anaphora are of special interest:

Conceptual pronouns and anaphoric island. Conceptual anaphora (Gernsbacher, 1991) are anaphora that don't have an explicit linguistic antecedent but one constructed from the context. For example, in a sequence of sentences like, "*I need a plate. Where do you keep them*?", "them" is not in agreement with the antecedent "plate". However, a sentence like "*where do you keep them*?" is more acceptable than sentences like "*where do you keep it*?". Generally, Gernsbacher (1991) found that this kind of anaphors was rated more natural and comprehended more quickly than the same sentence with singular pronoun if they refer to multiple item or events, generic types, or collective sets3. Thus, this type of anaphora refers to a special type of referent which can only be inferred (and that's why Gernsbacher calls them the "linguistically illegal" pronouns). Oakhill et al (1992) were interested in determining when the inference might be generated. They compared reading times of conceptual version like "I need a plate. *Where do you keep them*?" with an explicit plural version like "*I need some plates. Where do you keep them*?". Oakhill et al found that if the pronoun refers to generics or implied multiple items, conceptual items took more time to be read than the explicit plurals. However, in case of referring to collective sets, conceptual version needed shorter time to be read. Oakhill et al (1992) interpreted this longer reading time of conceptual pronouns in case of referring to generics or multiple items to mean that it is the area where the inferences are generated. So in response to "*where do you keep them*?", a set of plates can be inferred and introduced into the mental model. Accordingly, the anaphor here triggers the inference that makes the discourse coherent.

The other kind of indirect anaphora, anaphoric islands, is the reference into anaphoric clause to anaphoric island, as in 12 (from Garnham, 1999).

(12) *Jim reviewed that book and it will be published in the "linguistic inquiry."*The anaphor "it" here refers to the review even though it is not explicitly mentioned in the text .

To get a fuller picture about the effect of anaphoric referentiality on discourse comprehension, it might be appropriate to briefly present available facts about the time course of this effect. There is

considerable evidence supporting the immediacy hypothesis on anaphora resolution. In their review of the subject, Garrod and Sanford (1989) reviewed evidence for immediate pronoun resolution if the referent was the protagonist, in focus, or was introduced through a proper name. However, they also mentioned Ehrlich and Rayner (1983) study, which showed evidence for some delay in pronoun resolution especially for pronouns with distant or non focused antecedent. Accordingly, Garrod and Sanford (1989) proposed a "weak immediacy hypothesis" where they distinguished between the onset of an anaphoric resolution and its completion. The immediacy hypothesis seems to be also consistent with the results of an ERP[4] study (Van Berkum et al, 1999) that showed evidence for the beginning of searching for a unique antecedent 300 to 350 mesc after the offset of the NP. In an eyetracking[5] study, Garrod et al (1994) presented their subjects with short paragraphs about two characters but only one of them was always in the focus. The event in the target sentence could be congruent with previous context or not depending on the antecedent assigned to the anaphor. For example, after a story about a male life guard and a female inexperienced swimmer, subjects read sentences like 13.

(13) *Within seconds, she/he sank/jump into the pool*.

Conflict between the pronoun and the verb reflects pronoun resolution and might indicate where this process occurs. Garrod and Sanford (1989) found an immediate pronoun resolution as reflected in long first pass time [6] on the verb region, in case of pronoun-verb incongruency. However, this occurred only when the pronoun was referring unambiguously to a focused antecedent. In another experiment in the same study, they used a name and definite NP instead of the pronoun. They found that anaphoric resolution was delayed and appeared only on the second pass times7.

Thus, it seems that the timing of anaphoric resolution is not an all or none process. Rather, It might be safe to assume immediate resolution for the content free anaphoric devices like pronouns but not for proper names and definite NPs, which might need more time to be updated and integrated within the discourse representation .

A rather different picture is obtained for the time course of

temporal anaphora8. In an attempt to determine when temporal anaphora might be resolved, Dickey (1999) carried out a series of self-paced reading experiments. In these experiments, he had his subjects read unambiguous past tense sentences, like 14, that were preceded by either past-tense context (compatible context) or future-tense context (incompatible context).

(14)*He took/ the hope diamond/ and an emerald/ during his escape.*

The logic was that the timing of anaphoric resolution can be determined on the basis of detecting the clash between the sentence and the context in the incompatible context. This detection would be indicated by a significant difference between the sentences as presented in two context conditions. The results indicated that the only reliable difference was at segments 2-3, the segment at which the event description in the sentence is complete. This pattern of results was maintained both transitive and intransitive verbs (expt 2). Accordingly, it can be concluded that the time course for the temporal anaphora resolution is different from that of the pronoun anaphora. While pronoun resolution is immediate, the resolution does not seem to appear before the V+ direct object region, in case of temporal anaphora .

Thus, these developments in conceptualizing anaphora provide better understanding of its role in discourse coherence and comprehension. Of particular importance in this context is the transition from focusing on traditional anaphoric factors (e.g., distance between anaphora and antecedent, and similarity between them) to emphasizing the more conceptual concerns related to cognitive load and mental models associated with different forms and levels of anaphora .

3-Implicit Causality

Implicit causality is a property of transitive verbs where one of verb's arguments is the cause of the event or attitude expressed in the sentence. The term "implicit causality" was coined by Garvey and Caramazza (1974) to account for their subjects' performance on a sentence completion task. That is, when asked to continue sentences like "*John questioned Mary because…*", they will most likely ascribe the cause of the questioning in the sentence to the surface subject (the

noun phrase (NP) John) rather than the object NP (Mary). Accordingly, they will more likely continue it with phrases like "*because he wanted to know the truth*" instead of "because she was suspicious". Accordingly, the verb "to question" is considered an NP1 verb as it is biased toward first surface NP (grammatical subject) to be the cause of the event. On the other hand, a verb like "to praise" is considered an NP2 verb because it is biased toward considering the second surface NP (the grammatical object) the cause of the event. Therefore, subjects tended to continue sentences like "John praise Mary because…" with phrase like "*because she was successful*" rather than with phrases like "because he was in a good mood". Garvey and Caramazza (1974) claimed that verb's implicit causality bias is determined by what they called "locus of underlying cause". To support their point, Caramazza et al (1977) had their subjects read sentences with either NP1 verb or NP2 verb bias. The continuation was either consistent or inconsistent with the implicit causality bias. They found that subjects were faster in reading the sentence and in naming pronoun referent when the continuation was consistent with implicit causality bias than when it was not. For example, after an NP1 biased verb as in "Roy questioned Anthony..", subjects took less time to read "*because he wanted to learn the truth*" (consistent continuation) than to read "*because he had not told the truth*" (inconsistent continuation).

The effect of implicit causality on discourse processing was seen as a focusing device (e.g., Long & De Ley, 2000). It was claimed that implicit causality establishes the more prominent verb's argument (the cause of the event or the action) as the focus of attention. Being salient, the focused argument becomes more likely to be referred to later in the discourse. Moreover, pronouns that refer to this argument are easier to be resolved and their antecedents are faster to be recognized than unfocused argument.

McKoon et al (1993a) attributed the focusing effect of implicit causality to what they called initiating- reacting distinction. They argue that implicit causality is biased toward the argument that initiates the event or the action expressed in the sentence. That's why, according to them, the verb "to confess" is an NP1 verb as people

confess something they have already done. On the other hand, a verb like "to thank" is an NP2 verb because we thank others for something they have done for us. It should be noted that initiating-reacting distinction has nothing to do with the surface word order in the sentence (subject-verb-object in English). In other words, grammatical subject is the initiator while grammatical object is the reactor in NP1 verbs. In NP2 verbs, on the other hand, the grammatical object is the initiator and the grammatical subject is the reactor .

To test this hypothesis about attributing implicit causality's focusing effect to the initiating-reacting distinction, McKoon et al (1993b) investigated whether the character in the initiator role is more accessible than the character in reactor role as indicated by pronoun resolution. In a series of experiments, they presented their subjects with a short text. The last sentence included an implicit causality verb that was either biased toward NP1 as "*James infuriated Debbie because*... ", or NP2 as "*Diane valued Sam because*…". For each sentence, the continuation was either consistent or inconsistent with the verb bias. Accordingly, for sentences with NP1 verbs, subjects had continuation such as (a) "*he leaked important information to the press*" (consistent), or (b) "*she had to unite all the speeches*" (inconsistent). Similarly, for sentences with NP2, they had either consistent continuation (a) "*because he knew how to negotiate*", or inconsistent continuation (b) "*because she never knew how to negotiate*". Their logic was as follows. If implicit causality is biased toward the initiator, then the initiator will be active on reading the sentence continuation. Accordingly, NP1 verbs will activate the grammatical subject (the initiator "James" in the example above) on reading the continuation. On the other hand, NP2 verbs will activate the grammatical object (the initiator "Sam" in the example above) on reading the continuation. Therefore, it was predicted that recognizing the initiator's name will be faster after reading the consistent continuation where it was referent of the pronoun than after reading the inconsistent continuations where initiator's name was different from the referent of the pronoun. For example, in NP1 sentences, recognizing the grammatical subject "James" should be faster after the consistent continuation (where it is the referent) in comparison to its

recognition after the inconsistent continuation which has different antecedent for the pronoun (Debbie). This prediction was confirmed. The character in initiator role (subject in NP1 sentences and object in NP2 ones) was more accessible than the character in reactor role. This result was interpreted to mean that verb implicit causality made a character (the initiator) more accessible in the discourse model and more easily recognized as an antecedent for future pronoun. This, in turn, facilitates recognizing the character (initiator) as a test word. On the other hand, if the activated initiator character was not the antecedent of the pronoun (as in inconsistent continuation), this leads to more effort to activate the reactor as the referent of the pronoun and make the initiator less accessible .

McKoon et al (1993b) study presented above presents a new evidence for the role of implicit causality in establishing an NP at the focus of attention during building a mental model of the discourse. However, it is not very convincing in regard to presenting an evidence for their initiating- reacting distinction. The results can be attributed to the accessibility of the argument that fits with the implicit causality bias. In this case, these results are comparable to earlier findings (e.g., Caramazza, Grober, Garvey, and Yates, 1977) that continuation that is consistent with implicit causality bias are faster to be read and its argument is more likely to be a referent for future pronouns. Accordingly, an evidence for McKoon et al's initiating- reacting still needs to be presented. Again, here we can see that implicit causality's role in discourse processing is not related to the linguistic characteristics of the verb. Rather, this role can be attributed to the verb's contribution in building a mental model of the situation expressed in the discourse .

Another area of controversy is related to the time course of comprehending and using implicit causality. The issue is of particular importance in regard to getting a more complete picture of how implicit causality affects discourse processing. In general, there are two main proposals in regard to the timing of implicit causality information: the focus account and the integration account (see Garnham et al, 1996). Both accounts agree that the processor gets information about the implied cause and makes it active. This

information is used later to resolve pronoun with the implied cause as the likely antecedent. However, the focus account considers this process immediate as the implied cause becomes the most activated entity in reader's mental model and is used to resolve later pronouns as soon as they are encountered. On the other hand, according to integration account, using information about implied cause occurs later when the reader integrates the representations of two clauses into a single model .

No clear cut evidence is presented for any of the two accounts. Some researchers (e.g., Greene & McKoon, 1995; McDonald& MacWhinney, 1995) support the focusing effect whereas others (e.g., Garnham et al, 1996, Stewart et al, 2000) support the integration account. The discrepancy is rather surprising as the same task (probe latency task9) was used in most of these studies as explained below. Obviously, finding a congruency effect10 immediately after the pronoun supports the focusing account while finding an effect later or at the end of the sentence supports the integration account. In McDonald and MacWhinney's (1995) study, subjects listened to sentence as "*Beth disappointed Pam bitterly because she was so hard hearted at the anniversary party*". Probe words (NP1: Beth and NP2: Pam) was presented visually with different intervals: 100 ms after the second name, (2) immediately after the pronoun, (3) 200 ms after the pronoun, or (4) at the end of the sentence. They found that for NP1 sentences, NP1 probes were processed faster than NP2 on all intervals. This result was attributed to the first-mentioned effect (the name mentioned first in the sentence is verified faster). On the other hand, for NP2 sentences, reaction times to NP2 probe were faster at points (2) and (4) in comparison to reaction times at points (1) and (3). This result was interpreted to support the focusing account because of the facilitation that occurred after the pronoun in (2). Similar results were found in Greene and McKoon (1995) study. They compared reaction times to probes presented before and after the verb. They found that reaction time was faster for probes presented after the verb in comparison to probes presented before the verb.

On the other hand, Garnham et al (1996) gave different patterns of results. When the probe name appeared after the pronoun, data

showed first-mention effect regardless of the verb bias. When the probe name appeared at the end of the sentence, subjects responded faster to the probe consistent with verb bias showing congruency effect. Moreover, unsatisfied with the probe task, Stewart et al (2000) used self-paced reading technique11. As in previous research, they manipulated the verb bias and the congruency. Then critical manipulation, however, was in anaphora form when they used either proper name or pronoun as 15 and 16:

(15) Daniel apologized to Arnold because Daniel had been behaving selfishly .

(16) Daniel apologized to Arnold because He had been behaving selfishly.

The logic behind this manipulation was that proper names signal change in theme but pronouns refer to focused antecedents and are sensitive indicators of focus. Accordingly, if their referents are in focus, they will be read faster than the proper names (focus account prediction of interaction between reading time and anaphora form). On the other hand, if their referents are focused at the end of the sentence, no interaction should be found though congruent sentences should be read faster than incongruent sentences (integration account prediction of main effect with no interaction). The predictions of integration account were confirmed in experiment1 and maintained in experiment2 when they added a phrase at the end of the main phrase before the connective clauses. For more support of integration account, Stewart et al manipulated depth of processing by presenting shallow and deep questions after sentences in experiment3. Consistent with prediction of integration account, they found an interaction between congruency effect and level of processing. Congruency effect increased with deeper processing that supports more integration. In last experiment, they manipulated pronoun ambiguity. In some case the subordinate phrase included two verb arguments of the same gender, so the pronoun is ambiguous and can't differentiate between them. In some other cases, the pronoun was unambiguous as it agreed with one verb argument but not the other. Again, here also they got the same result: a main effect of congruency but no interaction between congruency and pronoun ambiguity.

The last study (Long & De Ley, 2000) presented in this regard tried to determine the condition in which each of focusing or integration accounts can be correct. Using probe recognition task, Long & De ley (2000) investigated the accessibility of the two verb arguments in their sentences. An important aspect of this study is that the researchers manipulated the reading skill as a between subject factor. They divided subjects into skilled readers and less skilled readers depending on Nelson- Denny Reading test. Results indicated that the data of the less skilled readers replicated the results of Garnham et al (1996) and Stewart et al (2000) which supported the integration account. That is, among this group, probe recognition was affected by implicit causality only at the end of the sentence and they showed first- mentioned effect. As for skilled readers, there was no first- mentioned effect and there was an effect of implicit causality in earlier region like after the pronoun. These results, however, were found only for the NP2 verbs. Long & De Ley's (2000) results indicate that different preferences to use either integration or focusing strategies may reflect individual differences in reading, and that readers' characteristics should be taken into account.

4-Syntactic Cues

The role of syntactic aspects of the sentence in discourse processing was theoretically motivated by the work of Givon (1989). According to Givon, grammatical cues function as mental processing instructions and can be thought of as focusing devices. As for discourse comprehension, the grammatical cues work as focusing devices that make an event or a referent more salient and consequently more likely to be mentioned later in the discourse. This, in turn, makes the referent more easily accessible as a referent for pronoun, which leads to a more coherent discourse .

In a series of experiments, Birch and her colleagues (Birch & Garnsey, 1995, Birch & Rayner, 1997, Birch et al, 2000) studied the effect of syntactic cues as focusing devices. In these experiments, they used two types of syntactic cues used to indicate focus in English: (1) The first cue was the <u>There insertion</u> as in "*There was this mugger who had attacked an elderly lady*", and (2) the <u>It- cleft</u> as in "*It was the mayor who refused to answer a reporter's question*". (examples

are from Birch et al, 2000, PP.288- 289).

Birch and Garnsey (1995) showed that, in general, memory for syntactically focused word was better than memory for their counterpart in non focus version of the same sentence. Accordingly, they compared focused vs. non-focused words, as in (17) and (18.(

(17) *It was the lion who stole the show at the circus.*

(18) The giraffe that the lion attacked couldn't run away.

They found that syntactically focused words were better recognized as seen before in a delayed recognition task12, where words were presented for recognition after 10 seconds interval, after reading the sentence, during which subjects counted backward (experiment 2). These focused words were also better named in a delayed naming task13 (experiment 3). However, difference between focused and non-focused words was not significant in regard to subjects' performance on immediate recognition task, where words were presented immediately after the sentence with no interval (experiment 1).

Studying the time course of focusing effect could substantiate the previous claim. Using eyetracking technique, Birch and Rayner (1997) compared reading times and eye movements during reading focused vs. non-focused words. They used the same syntactic focusing devices of There-insertion and It- cleft. Consistent with the results of off-line and probe recognition studies, they did not find significant differences between focused and non-focused words on measures of initial processing like first pass and gaze duration14. Rather, they found differences between them on reprocessing measures; namely second pass14 and regression.

In a latter study, Birch et al (2000) replicated the major trends of previous studies (e.g., Birch & Garnsey (1995) and Birch & Rayner (1997)). Using a continuation paradigm15, they found that subjects referred to focused words more than non focused ones in their continuation of the last sentence of a short story. Two main aspects of this series of experiments are worth mentioning: First, the material was a stretched written discourse, which gave them the chance to study the effects of syntactic focusing over a wider range than the sentences used in Birch and Garn sey (1995) study. The second new aspect was related to the control sentence. They noticed that control

sentences in Birch and Garnsey (1995) study de-emphasized the focused word. They reasoned that the obtained focusing effect could be attributed to either the existence of the focusing cue or to the inhibition of de-emphasized word. Accordingly, they used "neutral words" to avoid this confounding effect. Accordingly, for the experimental sentence "*there was this mugger who had attacked an elderly lady*", the deemphasizing control sentence (as used in Birch & Garnsey 1995) was "*The police caught a mugger who attacked an elderly lady*". On the other hand, the neutral control sentence (as used in Birch et al, 2000) would be "*A mugger had attacked an elderly lady*". It can be seen that in the neutral condition, the non-focused word was presented as a new entity, though not focused. This was assumed to give a more refined method of comparison between reaction times for focused and non-focused words. However, as in experiment 1 in Birch and Garnsey (1995), Birch et al (2000) found no significant difference between recognition times for focused and non-focused word as measured by immediate recognition task (experiment 2A and 2B). Again, subjects were significantly faster in recognizing focused words than non-focused words in delayed recognition task (experiment 4).

From previous experiments, it can be seen that the effect of syntactic focusing is not immediate. Rather, it occurs in a later stage. Birch et al (2000) interpreted that in terms of different levels of activation of the words in working Memory. They argue that in case of immediate recognition, focused and non-focused words are still active in the working memory and it is reasonable not to expect difference in availability between them. On the other hand, in case of delayed recognition, focusing helps to keep the focused word in an active state for longer time and, consequently, to be easily integrated in further discourse representation.

Combined together, these studies show that the effect of syntactic focusing is not related to facilitating initial lexical recognition. Rather, it is related to providing the focused word with more accessibility and deeper level of processing that facilitate its integration with more permanent mental representation of the discourse .

The role of syntactic cues and syntactic focusing began to be

recognized even among bottom-up theorists who used to ignore syntactic variables like Kintsch. To account for syntactic cues effects as described above, Kintsch (1992) studied the effect of focusing an NP with the topicality cue "this" on the performance of his Construction-Integration (C-I) model16. More specifically, he compared matrices of connection strengths resulting from simulating his model's processing of two versions of the text. The only difference between the two versions was that one of them included an NP "egg" presented with an indefinite article "an", while the same NP was focused by "this" in the second version. The focused NP's self connection strengths were increased from 1 to 2 to assure its centrality in the discourse. This manipulation led to different memory representation networks for the two versions with low correlation between them .

Kintsch's (1992) attempt to include the role of grammatical cues in his C-I model is not very convincing. Methodologically, the difference between representation matrices of the two versions of the read text is not surprising, taking into account the a priori decision to double the self connection strengths of the focused NP but not for the non focused one. Accordingly, the critical question might be whether Kintsch's model that depends on an initial unselectivity (the construction process) has a room for such modification. Kintsch did not give any theoretical reason for his specific change in the self connection strength of the focused NP, or for not including any other change in it. In this case, the change that Kintsch made may not be more than a guess. More work, accordingly, needs to be done to specify the role of syntactic cues in Kintsch's model, and other bottom up models, if the aim is more than generating a better simulation matrix.

5-Linguistic Markers

Linguistic markers are a group of expressions that don't constitute a single, well-defined grammatical class but comprise a functionally related group of items drawn from the other classes. They are generally assumed to convey information about the discourse structure and the relationships between its parts.

Linguistic markers are mentioned in the literature by a variety of

titles such as linguistic connectives, linguistic devices, discourse markers, cue phrases, text-signaling devices, surface markers, and segmentation markers.

Despite some disagreement on their classification (e.g., Caron, 1997; Gernsbacher, 1997; Schourup, 1999; Townsend, 1983, 1997), most researchers agree that linguistic markers can be divided into the four following categories:

1. Additive markers (and, or)
2. Causal markers (because, since, so, consequently, therefore, ..etc(
3. Adversative markers (although, while, despite, but, whereas, however, ..etc)
4. Temporal markers (after, before, simultaneously, until, afterward, ..etc).

Studying linguistic markers began in 1980s as part of the interest in the effect of structural, surface characteristics of the text on its processing. The first two studies in this regard were by Britton and his colleagues (Britton, Glynn, Mayer, & Penland, 1982), and by Haberlandt (1982) .

In Britton et al study, they compared their subjects' memorization and comprehension of two equivalent versions of sentences with or without signaling connectives. Both versions were presented within texts like Breeder reactors produce more nuclear fuel than they consume. *In addition*, these reactors would operate without adding noxious combustion products to the air. *It is in the light of these considerations that U.S. atomic Energy Commission ….* Obviously, the without signaling version of the text didn't include the underlined connectives. They found that there is no effect of connectives on free recall measure, but texts with connectives required significantly less inspection time (time needed to read one page) than texts without connectives. More importantly, the time needed to respond to a secondary task (releasing a key on hearing a click, a task that was assumed to consume some of the cognitive capacity needed for reading) was less if this task was accompanied by reading a connected text than unconnected one. This result was interpreted, in accordance with the literature on secondary task paradigm, to mean that connectives made reading easier and, consequently, the reader could

allocate more resources to the secondary task. In other words, texts with linguistic markers or connectives require less cognitive capacity to be processed than texts without linguistic connectives .

Unlike Britton et al study that measured the effect of markers on the text as a whole without indicating a specific segment on which the effect is predicted to appear, Haberlandt (1982) was interested in two specific phrases. He presented his target sentences in stories like

The jet had just taken off.

The left engine caught fire.

The passengers were terrified.

They thought the plane would crash.

Target *However, the pilot made a safe landing vs .*

The pilot made a safe landing.

Haberlandt compared subjects' performance on sentences with causal or adversative connectives with their performance on the sentences with no connectives at all. In both conditions, he was interested in comparing the performance on (1) the phrase which follows the connective (first phrase in case of no connective), (2) the second phrase that is further separated from the connective toward the end of the sentence, and (3) the entire sentence. Haberlandt found that sentences with connectives were read significantly faster than sentences without connectives. The same effect was found in the first phrase that follows the connective. However, the effect was not significant for the late phrase or for the difference between adversative and causal connectives.

The effect of linguistic markers was confirmed again by Sanders and Noordman (2000) who found faster reading time and verification latencies of sentences with linguistic markers than the sentences without them. However, linguistic markers didn't increase the amount of recalled information.

Markers' effect was also found in understanding spontaneous talk during conversation. In a study by Fox Tree and Schrock (1999), they compared the availability of words (using word monitoring task17) after Oh to their availability with the Oh either replaced by a pause or removed entirely. They found that the recognition of words was faster after Oh than when Oh either replaced by a pause or excised entirely.

Again, these results were interpreted to mean that Oh was used to integrate the information in spontaneous talk.

In a series of experiments, Townsend (1983) addressed the issue of the differences between different types of connectives. He used a variety of measures like synonymy judgement task (where the participant has to judge whether two words are synonymous or not), word naming task, recall rate, and reading time. In experiment I, he showed that the accessibility to the meaning and word recognition time of words in initial clauses was faster after the whole sentence than it was during listening to the first clause, in case of using "*while*" connective. This effect didn't exist for the connective "*since*". Again, in experiment II, response time to questions about active main clauses were faster than their counterpart questions about passive clauses in case of using "*since*" to mark the clause but not when the clause is marked by "*though*". In experiment IV, subjects' accessibility to the meaning of the final clause after reading a story was slower for the clauses introduced by non-causal connectives (e.g., although, before) than it was for clauses introduced by causal ones. In experiment VII, the time necessary to create a continuation to a two-sentence texts varied depending on the type of connective used to introduce the second sentence. The order was *therefore* < *afterwards* < *no connective* < *previously* < *meanwhile* < *however.* Generally, this indicates, Townsend argues, that causal and temporal connectives have more integrative effect than adversative ones .

This emphasis on integrative aspects of linguistic markers, especially the causal one, was crystallized by Millis and Just (1994) Connective Integrative Model which works through their proposed Delayed Reactivation hypothesis. According to this model, the reader constructs a representation of the first clause. When he encounters the connective, he knows that he must integrate previous clause with the incoming one. He proceeds to the second clause and represents it in working memory. By the end of the sentence, the reader reactivates the representation of the first clause and integrates it with representation of the second clause in a single representation. This process is assumed to cost time .

In their experiment, they compared the reading time, probe time,

and accuracy of comprehending sentences with inter-clause connective "*because*" that connects two clauses in a single sentence vs. the performance on the same two clauses presented separately without connective as two sentences. The example they mentioned in their paper is *The elderly parents toasted their only daughter at the dinner because Jill had passed the exams in the prestigious university* (inter clausal connection) vs. *The elderly parents toasted their only daughter at the dinner. Jill had passed the exams in the prestigious university* (separately presented sentences). They found that reading time was faster for the phrases following the connective than for the same phrases without connective. However, the reading time of the last word was slower in case of connective than in cases where there was no connective between the two sentences. These results were interpreted to mean that the connective facilitated realizing the relationship in the words that follow the connective. On the other hand, the integrative process occurs as the result of reactivating the representation of the first clause at the end of the entire sentence resulting in longer reading time of the last word. To support their claim, they compared the probe time to the verb from the first clause in two locations: one word after the connective and immediately after the last word of the second clause. They found that the presence of the connective increased the probe time when the probed word occurred earlier in the sentence than when it occurred after the last word in the sentence. This result was interpreted to mean that the entire sentence is fully activated only at the end of the sentence.

On the other hand, the causal connective "because" was found to play a significant role in processing different types of causal relations. Two different research groups (Noordman & de Blijzer, 2000 and Traxler, Sanford, Aked, & Moxey, 1997) found that evidential causal statements (as in 19) are more difficult to process than equivalent factual statement (as in 20).

(19)Because most distinguished students got bad grades, the teacher made some mistakes in evaluating his students' papers.

(20)Because he got tired after a long semester, the teacher made some mistakes in evaluating his students' papers.

Moreover, it was found that epistemic modals (such as I think or

must) would eliminate the difficulty associated with evidential causal statements, but not the factual ones (as in 21 and 22).

(21)Because most distinguished students got bad grades, I think the teacher made some mistakes in evaluating his students' papers.

(22)Because he got tired after a long semester, I think the teacher made some mistakes in evaluating his students' papers.

Mohamed and Clifton (2008) replicated these findings and introduced a third type of causal relations; the deductive causal statement (see also Mohamed, 2003) as in (23) below. They theorized that while factual statements represent a sequence of actual events, where event A (e.g., the teacher's being tired) causes event B (e.g., making mistakes in grading students' papers), evidential statements express a sequence of mental reasoning, in which event A (the failure of the distinguished students) is a reason/evidence to believe that B (the teacher's mistakes in grading the papers) is the case. And that is why, it is assumed, using epistemic modals helps the processor to understand that the statement is expressing a mental reasoning process rather than actual sequence of events. The deductive causal statements was introduced by Mohamed and Clifton to represent a case in which B (the teacher's mistakes in grading the papers) is implied in and deducted from the general/atemporal statement A (grading papers is a subjective process).

(23)Because grading a paper is a subjective process, the teacher made some mistakes in evaluating his students' papers.

Mohamed and Clifton classified both evidential and deductive causal statements as inferential or reasoning processes. While factual and evidential statements move from specific evidence or a reason to reach an event, the deductive statements move from a general rule or premise to reach a specific event. Because of this difference between specific and general subordinate clauses in evidential and deductive statements, respectively, it was predicted that deductive causal statements would be easier to process than evidential ones and more difficult than factual ones. Actual results supported these predictions. Moreover, it was found that epistemic modals have same facilitative effect for deductive statements, as in the case of evidential ones .

Moreover, Mohamed and Clifton found that processing different

types of causal relations is also affected by the verb type used in the causal statement. That is, they distinguished between two types of verbs: Psychological verbs (e.g., likes or feel) and action verbs (e.g., destroy or move). Mohamed and Clifton based their distinction on the notion of observability, and used the following criteria to distinguish between action and psychological verbs (from Mohamed & Clifton, 2008, p.42)

1. Action verbs are external and denote events that occur in the objective world, whereas psychological state verbs are internal and occur on the psychological or mental level .
2. Action verbs express acts that can be seen and described by everyone who has access to the situation, whereas psychological verbs express personal experience felt and evaluated only by the agent.
3. Action verbs express action that tends to have a delimited time with a relatively specific beginning and end, whereas psychological verbs tend to be relatively enduring with no specific beginning and end.
4. Indicating or describing a psychological verb requires an inference generation process, whereas describing an action verb can be direct .

Based on acceptability rates of different types of causal statements, Mohamed and Clifton found that using psychological verbs to express evidential causal relations makes these statements more acceptable and easier to process than using action verbs, whereas using these psychological verbs has negative effect on processing factual statements. On the other hand, using action verbs to express factual relations makes these causal statements more acceptable and understandable than using psychological verbs, whereas these action verbs has negative effect on processing evidential relations.

These results were attributed to the fact that psychological verbs express internal, hypothesized events, which are more consistent with the inferential nature of the evidential causal statements, while action verbs express concrete, observable events, which are more consistent with realistic nature of the factual causal statements .

In another line of research, the main focus is on the temporal

connectives and their functions. With the situational models of discourse, it is assumed that the reader tends to temporally locate each incoming event after the most recent one in the model. Accordingly, it has been proposed that readers assume by default that the order of reported events corresponds to the chronological order of them. This assumption was proposed by Dowty (1986) as the *temporal discourse interpretation principle* (TDIP) and is currently referred to as the Iconicity Principle (for a review of the origins of this principle, see Zwaan, 1996).

Consistent with this logic, Mandler (1986) found that mismatch between chronological order and temporal order increases sentence reading time. Using comprehension task18, Ohtsuka and Brewer (1992) found that with more deviation between the chronological order and the reported order, there was more difficulty in understanding the text. Similarly, Zwaan (1996) found that sentence reading time was slower in case of narrative time shift *(e.g., an hour later in Jamie turned on his PC and started typing. An hour later, the telephone rang)* than in case there is no such shifts *(e.g., a moment later in Jamie turned on his PC and started typing. A moment later, the telephone rang)*. Probe recognition time was longer for information about the previous event when it was followed by a time shift than when it was not .

Within the same line of research, Bestgen and Vonk (1995) found that probe recognition time for the contents of the sentences preceding the critical sentence with connective was longer in case of using temporal connectives like *then or around twelve thirty*, than in case of using no connective (experiment 1) or using additive connective like and (experiments 2 and 3). On the other hand, they could not find a significant difference in reading time between the marked and unmarked sentences .

Both Zwaan (1996) and Bestgen and Vonk (1995) interpreted their results within Gernsbacher's Structure Building Framework (Gernsbacher, 1990). In this model, discourse processing depends on three processes: using the first words or sentences to *lay the foundations* for the first substructure. If the new information is coherent with this initial structure, *it is mapped onto it*. If not, the

reader *shifts and initiates* a new structure. According to this logic, in case of the need to shift and initiate a new structure, the reader concentrates more on building the new structure (which costs time and cognitive effort) and, accordingly, previous information becomes less accessible. This was confirmed by results from Zwaan (1996) and Bestgen and Vonk (1995). However, while Zwaan found that time shift condition led to longer reading time than no shift condition, Bestgen and Vonk didn't find this effect in case of temporal connectives on the reading time. Zwaan (1996) attributed this discrepancy between his results and Bestgen and Vonk's results to the increasing memory load necessary for laying a new foundation. In his words, "the reader takes the time shift as a cue to decrease the activation of the information preceding the time shift and to set up a new time interval. This causes a momentary increase in on-line processing load." (Zwaan, 1996, p.1205). This logic was further supported by the results reported by Bestgen and Vonk (2000). They compared the reading time of highly congruent sentences (continuous texts) vs. weakly congruent sentences (discontinuous texts). They found that discontinuous texts were read slower than the continuous texts, but this effect disappeared with the usage of temporal connectives. Unlike Zwaan's findings, these results meant that temporal connectives (as topic shift devices) speeded the reading time of the marked (connected) sentences. Bestgen and Vonk (2000) interpreted their results within Gernsbacher's model to mean that temporal connectives work as segmentation markers that help the processor to abandon the nextness principle and to expect a topic shift. In other words, the temporal connectives help the processor to avoid mapping the sentence into the previous structure and to initiate building a new one. To support their point, they compared reading times of sentences with a temporal adverbial (around 11 O'clock) versus reading time of the sentences with a sentence adverbial (as usual) in experiments 3 and 4. They showed that temporal adverbial facilitated the topic shift while the sentence adverbial didn't.

Thus, to sum up this review, the general finding from previous research is that using linguistic markers (1) reduce the reading time of the sentence or clause that follows it, (2) improve the performance on

comprehension questions, though not on recall, and (3) increase the availability of previously stored information in case of continuity markers and reduces this availability in case of discontinuity, topic shift markers. These effects can have a major impact on discourse processing as they facilitate reading a continuous texts and make topic shift in discontinuous texts an easy and smooth Process. The theoretical interpretations presented for the markers' effects highlighted their integrative function within two theoretical frameworks:

1. Gernsbacher's Structure Building Framework. Avoiding the Nextness Principle, as applied to inter-sentence temporal connectives (e.g., Bestgen & Vonk, 1995; 2000, and Zwaan,1996), is assumed to override the default temporal correspondence between reported order and chronological order (Iconicity Assumption, Zwaan, 1996; also Mander, 1986)
2. Millis and Just's (1994) Connective Integration Model and Delayed Reactivation Hypothesis (mainly applied to inter-clause causal connective) and results in late integration of the clauses' mental representations.

However, the previous research and their interpretation have some problems that need to be addressed. Methodologically, different researchers measured the reading time of different region(s). While some were accurate and measured specific regions (e.g., Haberlandt, 1982; Millis & Just, 1994), most researchers measured the reading time of the whole sentence. The same problem can be found in relation to the probe recognition .

Another problem was the control sentences that were not identical to the experimental sentences in case of inter-clause connectives. For example, in Millis & Just (1994), they compare sentences with connectives like *John got a good grade because he studied hard* with the two sentences *John got a good grade and He studied hard*. Comparing reading time of a sentence with the reading times of two sentences can be hardly justified with respect to the integrative processes that occur once in the first case but twice in the second case. From the theoretical point of view, the major problem with the proposed theoretical interpretations is the vague and indeterminate

nature of the hypothesized integrative process. While some researchers consider it a reactivation of previous information (as measured mainly by probe recognition task), other researchers consider it a process of combining the representations of two sentences or two clauses. However, the question of how we can account for the integrative effect is still not fully answered .

6-Conclusion

It can be seen from previous linguistic analyses and empirical evidence that the sentential variables discussed in this study are critical to discourse representation. Mainly, the reader uses them to set up this mental representation and keep updating it as the text unfolds. Linguistic sentential variables play different but integrative roles in this regard. Syntactic cues helps to set up the focus of the text, the event, or character to be tracked in the following parts of the text. Anaphoric referentiality gives a sense of continuity in the text through maintaining the identity of objects, events, and characters. It helps the reader to avoid reading repeated information and determines what is given and what is new in the text and keeps track of changes in the status of this information (e.g., what is new in the current sentence is given in the following one). Implicit causality helps to indicate the initiators and reactors in the text. Initiators are usually assigned more central role in the text and to be focused on and followed by the reader. Also, change in the status of an initiator (to be reactor) or vice versa is an important topic shift that changes the text representation. Linguistic markers indicate the logical and temporal relations among parts of the text. They also can be the major indication of topic continuity or major change in the text.

Through sentential variables, the reader can have a coherent and dynamic representation of the text through following and updating the major tracks of its structure. Discrepancies among these variables enable the reader to detect major contradictions in the text. Also, sensitivity to these variables reduces the load necessary to maintain the text in memory by freeing the memory from keeping redundant, unnecessary details .

In addition to the psycholinguistic significance of these sentential variables, studying them has some applied implications. Sensitivity to

these variables may determine individual differences among readers. Training readers, particularly students, to be more sensitive to these variables may help in improving their reading skills. Also, sensitivity to these variables among authors and textbook writers can result in better organized and more comprehensible textbooks.

7-Endnotes

1. Relative clause is a clause that modifies a noun and includes a pronoun or other element which refers to this noun. Accordingly, in the sentence the man who came.., the relative pronoun who modifies the noun the man.
2. Ellipsis is the omission of one or more elements from a construction, especially when the omitted parts can be inferred from the context. For example if A asks "have you graded the assignment", B might answer elliptically "I have not", with the rest of the construction (graded the assignment" to be inferred .
3. Collective set is a noun that refers to individuals as a group. For example the word clergy is used to mean clergymen in general .
4. ERP is an abbreviation of event related potentials, which is a measure of electric activity in the brain. Using multiple electrodes to measure this activity in different brain regions results in getting graphs of brainwaves, and defines when and where electric activity moves over the surface of the cerebral cortex. ERP is used in psycholinguistic research to define areas responsible for understanding syntactic and semantic anomalies .
5. Eye tracking measure is measuring eye movements as person reads linguistic material (e.g., a word, sentence, or a short paragraph) or a sees a scene using an equipment called the eyetracker. It is commonly used in psycholinguistic research to determine how much time, in milliseconds, each subject spends in reading each segment (number and durations of fixations), in moving from one part to another (jumps or "saccades"), or in getting back to a part of the material that was read before (regression).
6. First pass time is the time spent or the sum of all fixations beginning with the reader's first fixation in region until the reader's gaze leaves the region. Usually, it is taken as a measure of initial processing of a region .
7. Second pass time is the time spent or the number of fixations in a region after leaving the region (or after the initial skip of it) and coming back to it. Usually, it is taken as a measure of late processing of a region.
8. Temporal anaphora is the interpretation of the reference time of a sentence depending on the reference time of the previous sentence. Accordingly, temporal anaphora is a tool to coordinate the time frames in two sentences. That might apply to pronouns referring to back to noun phrases in a previous sentence (e.g., the mail arrived this morning. I was at home then), and to the time introduced by the tense of a sentence (Ali woke up. It was cold then).
9. Probe latency task is an experimental paradigm in which subject is

presented with a sentence or group of sentences to read. Later on, he is presented with a word (a probe) and he has to decide as quickly and accurately as possible whether this probe is part of the original material. Accuracy and latency (time between presenting the probe and the reader's response) is taken as measure of how active (in the working memory) each part of the sentence is.

10. Congruency effect is the (usually) facilitative effect of the consistency between the verb's implicit causality bias in a sentence and the cause-effect relationship in the following sentence .
11. Self paced technique is an experimental paradigm common in experimental psycholinguistics. In this paradigm, the subject is presented by a linguistic material and is asked to read it. The material, though, is presented in segments, and he controls the presentation of these segments by pressing a key that results in removing the sentence that was already read and presenting the following one. Time between presenting a segment and removing it is measured in milliseconds and is considered a function of comprehension .
12. Delayed recognition task is task in a group of sentences or separate words are presented on a computer screen for a short period of time (e.g., 5 seconds). Then, subjects are shown an interfering stimulus (e.g., a picture) before the recognition task starts. In this task , a series of words are shown and subjects have to decide as accurately and quickly as possible, for every word whether it appeared in the previously presented sentences or not. This task is different from the immediate recognition task where the recognition task starts immediately after presenting the sentences with no interfering stimuli .
13. Delayed naming task is task in which a group of sentences are presented on a computer screen followed by a group of interfering stimuli (as in the delayed recognition task; see above). Then, a group of words are presented individually on the screen for up to 1 second, and subjects have to pronounce every word into a microphone as accurately and quickly as possible .
14. Gaze duration is equivalent to the first pass time (see above) if the region is defined to be one-word unit .
15. Continuation paradigm is task in which an incomplete sentence is presented after reading a sentence or a short paragraph. In this task, subjects have to complete this sentence, either in writing or by choosing between different possible continuations, based on their understanding of the previously read sentence or paragraph .
16. C-I (construction- integration) model is model of text processing that is

proposed by Walter Kintsch, a professor at University of Colorado in the U.S. According to this model, there are two stages of text processing. The first is the construction (C) of low level, automatic representation of the encountered sentence in the text, and this representation is based on the physical characteristics of the this sentence (e.g., length of the sentence, frequency of words in the sentence, the distances between anaphora and antecedents, etc…). The second stage of this model is integrating the representations of different sentences in high level, meaning-based representation .

17. Word monitoring task is a task in which subjects are required to listen to a set of sentences, and to keep track of a certain word in each sentence. Subjects have to press a certain button on hearing the tracked words .
18. Comprehension paradigm is a task in which a question is presented after each online read sentence (that is presented on a computer screen and reading time is measured). The subjects have to answer this question (or to choose the correct answer) based on their understanding of the sentence that was read before.

8-References

Almor, A. (1999). Noun-phrase Anaphora and Focus: The Information Load Hypothesis. *Psychological Review*, 106, 748-765.

Bestgen, Y., & Vonk, W. (1995). The Role of Temporal Segmentation Markers in Discourse Processing. *Discourse Processes*, 25, 385-406.

Bestgen, Y., & Vonk, W. (2000). Temporal Adverbials as Segmentation Markers in Discourse Comprehension. *Journal of Memory and Language*, 42, 74-87.

Birch, S., Albrecht, J., & Myers, J. (2000). Syntactic Focusing Structures Influences Discourse Processing. *Discourse Processes*, 30, 285-304.

Birch, S., & Garnsey, S. (1995). The Effect of Focus on Memory for Words in Sentences. *Journal of memory and Language*, 34, 232-267.

Birch, S., & Rayner, K. (1997). Linguistic Focus Affects Eye Movements During Reading. Memory and Cognition. 25, 653-660 .

Britton, B. K., Glynn, S. M., Mayer, B. J. F, & Penland, M. J. (1982). Effects of Text Structure on Use of Cognitive Capacity During Reading. *Journal of Educational Psychology*, 74, 51-61 .

Caron, J. (1997). Toward a Procedural Approach of the Meaning of Connectives. In J .Costermans & M. Fayol (Eds.), *Processing Interclausal Relationships: Studies in the Production and Comprehension of Text*, (pp.53-74). Lawrance Erlbaum Associates Inc.

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of Anaphoric Pronouns. *Journal of verbal learning and verbal behavior*, 16, 601-609.

Chater, N., & Christiansen, M. (1999). Connectionism and Natural Language Processing In S. Garrod and M. Pickering (Eds.), *Language Processing* (pp. 233-279) .(Psychology Press Ltd, publishers.

Dickey, M. (2000). *Processing of tense*. Ph.D. dissertation. UMass, Amherst.

Dowty, D. R. (1986). The Effects of Aspectual Class on the Temporal Structure of Discourse: Semantics or Pragmatics? *Linguistics and philosophy*, 9, 37-61.

Ehrlich, K., & Rayner, K. (1983). Pronoun Assignment and Semantic Integration During Reading: Eye Movements and Immediacy of Processing. *Journal of Verbal learning and Verbal Behavior*, 22, 75-87.

Fauconnier, G., Turner, M. (1998). Conceptual integration networks. *Cognitive Science*.187-133 ،22 ،

Fox Tree, J. E. & Schrock, J. C. (1999). Discourse Markers in Spontaneous Speech: Oh What a Difference an Oh Makes. *Journal of Memory and Language*, 40, 280-295.

Garnham, A. (1999). References and anphora. In S. Garrod and M. Pickering (Eds.), *Language processing*. Psychology Press Ltd, publishers.

Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. (1996). The Locus of Implicit Causality Effects in Comprehension. *Journal of Memory and Language*, 35, 517.543-

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The Role of Different Types of Anaphor in the On-Line Resolution of Sentences in A Discourse. Journal of Memory and Lnaguage, 33, 39-68.

Garrod, S., & Sanford, A. (1989). Discourse Models as Interface Between Language and the Spatial World. *Journal of Semantics*, 6, 147-160 .

Garvey, C., & Caramazza, A. (1974). Implicit Causality in Verbs. *Linguistic Inquiry*, 5464-459 ،

Gernsbacher, M. A. (1990). *Language Comprehension as Structure* Building. Hillsdale ،NY: Erlbaum.

Gernsbacher, M. A. (1991). Comprehending Conceptual Anaphors. Language and Cognitive Processes, 6, 81-105 .

Gernsbacher, M. A. (1997). Coherence Cues Mapping During Comprehension. In J .Costermans & M. Fayol (Eds.), *Processing Interclausal Relationships: Studies in the Production and Comprehension of Text*, (pp. 3-21). Lawrance Erlbaum Associates.

Gibson, E. (1998). Linguistic Complexity: Locality of Syntactic

Dependencies. *Cognition*, 68, 1-76 .

Givon, T. (1989). *Mind, Code, and Context: Essays in Pragmatics*. Hillsdale NJ: Lawrance Erlbaum Associates .

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, Names, and the Centering of Attention in Discourse. *Cognitive Science*, 17, 311-347.

Gordon, P.C., & Chan, D. (1995). Pronouns, Passives, and Discourse Coherence. *Journal of Memory and Language,* 34, 216-231.

Greene, S., & McKoon, G. (1995). Telling Something We Can't Know: Experimental Approaches to Verbs Exhibiting Implicit Causality. *Psychological Science*, 5, 262- 270

Haberlandt, K. (1982). Reader Expectations in Text Comprehension. In J. F. Le Ny & W. Kintsch (Eds.), *Language Comprehension*, (pp. 239-249). Amsterdam: North Holland.

Hankamer, J., & Sag, I. (1976). Deep and Surface Anaphora. *Linguistic Inquiry*, 7, 391 - .428

Haviland, S., & Clark, H. (1974). What's New? Acquiring New Information as A Process of Comprehension. Journal of Verbal Leaning and Verbal Behavior, 13, 513-521.

Heim, I. R. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D .dissertation. UMass, Amherst. GLSA .

Johnson-Laird, P. N. (1983). *Mental Models: Towards a cognitive science of language. inference, and consciousness*. Harvard University Press.

Kintsch, W. (1992). How readers construct situation models for stories: The role of syntactic cues and causal inferences. In A. F. healy, S. M. Kosslyn, & R. M .Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William Estes* (pp. 135-155). London: Lawrance Erlbaum Associates ،Ltd.

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition* Cambridge University Press .

Kintsch, W., & Van Dijk, T. A. (1978). Toward a Model of Text Comprehension and Production. Psychological Review, 85, 363-394.

Long, D., & De Ley, L. (2000). Understanding Anaphors in Story

Dialogue. *Memory*, 28 ، 738-731

Mandler, J. M. (1986). On the Comprehension of Temporal Order. *Language and Cognitive Processes*, 1, 309-320.

Mauner, G., Tannenhaus, M., & Carlson, G. (1995). A Note on Parallelism Effects in Processing Deep and Surface Verb-Phrase Anaphora. *Language and Cognitive Processes*, 10, 1-12 .

McDonald, J., MacWhinney, B. (1995). The Time Course of Anaphoric Resolution :Effects of Implicit Verb Causality and Gender. *Journal of Memory and Language*, .566-543 ،34

McKoon, G., Greene, S., & Ratcliff, R. (1993a). Discourse Models, Pronoun Resolution, and the Explicit Causality of Verbs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 82-91 .

McKoon, G., Ward, G., Ratcliff, R., & Sproat, R. (1993b). Morphosyntactic and Pragmatic Factors Affecting the Accessibility of Discourse Entities. *Journal of Memory and Language*, 32, 56-75 .

Millis, K. K., & Just, M. A. (1994). The Influence of Connectives on Sentence Comprehension. Journal of Memory and Language, 33, 1280-147.

Mohamed, M. T. (2003). Deductive Causal Relations. In Luis Alonso-Ovalle (Ed.), On *Semantic Processing* (PP. 167-187). Amherst, MA: GLSA.

Mohamed, M. T. (2007). Psycholinguistic Models of Text Processing: Discourse Analysis from a Cognitive Perspective. *Annals of Arts and Social Sciences Monographs*. Vol. 26, Monograph 261 .

Mohamed, M. T., & Clifton, C. (2008). Processing Inferential Causal Statements :Theoretical Refinements and the Role of Verb Type. *Discourse Processes*, 45, 24- 51

Murphy, G. L. (1985). Processes of Understanding Anaphora. *Journal of Memory and Language*, 24, 290-303.

Noordman, L., & de Blijzer, F. (2000). On the Processing of Causal Relations. In F .Couper-Kuhlen & B. Kortman (Eds.), *Cause, Condition, Concession, Contrast* .Berlin: de Gruyter .

Oakhill, J., Garnham, A., Gernsbacher, M., & Cain, K. (1992). How Natural Are Conceptual Anaphors? *Language and Cognitive Processes*, 7, 257-280.

Ohtsuka, K., & Brewer, W. F. (1992). Discourse Organization in the Comprehension of Temporal Order in Narrative Texts. *Discourse Processes*, 15, 317-336 .

Sag, I., & Hankamer, J. (1984). Toward a Theory of Anaphoric Processing. *Linguistics and Philosophy*, 7, 325-345.

Sanders, T. M., & Noordman, L. G. M. (2000). The Role of Coherence Relations and Their Linguistic Markers in Text Processing. *Discourse Processes*, 29, 37-60.

Schourup, L. (1999). Discourse Markers. *Lingua*, 107, 227-265.

Stewart, A., Pickering, M., Sanford, A. (2000). The Time Course of the Influence of Implicit Causality Information: Focusing Versus Integration Accounts. *Journal of Memory and Language*, 42, 423-443.

Tannenhaus, M. K., & Carlson, G. N. (1990) Comprehension of Deep and Surface Verb Phrase Anaphors. *Journal of Memory and Cognitive Processes*. 5, 257-280.

Townsend, D. J. (1983). Thematic Processing in Sentences and Texts. *Cognition*, 13 ، 261-223

Townsend, D. J. (1997). Processing Clauses and Their Relationships During Comprehension. In J. Costermans & M. Fayol (Eds.), *Processing Interclausal Relationships: Studies in the Production and Comprehension of Text*, (pp.265- 282 Lawrance Erlbaum Associates, Inc

Traxler, M., Sanford, A., Aked, J., & Moxey, L. (1997). Processing Causal and Diagnostic Statements in Discourse. Journal of *Experimental Psychology: Learning, Memory, and Cognition*. 23, 88-101.

Van Berkum, J., Brown, C., & Hagoort, P. (1999). Early Referential Context Effects in Sentence Processing: Evidence from Event-Related Potentials. *Journal of Memory and Language*, 41, 147-182.

Warren, T., & Gibson, E. (2000). Effects of Discourse Status on Reading Times :Implications for Quantifying Distance in a

Locality Based Theory of Linguistic Complexity. Poster Presented at the 13th annual *CUNY conference on Human Sentence Processing*. La Jolla, CA .

Warren, T., & Gibson, E. (2002). The Influence of Referential Processing on Sentence Complexity. *Cognition*, 85, 79-112 .

Zwaan, R. A. (1996). Processing Narrative Time Shift. Journal of *Experimental Psychology*: *Learning, Memory, and Cognition,* 22, 1196-1207.