Tel Aviv University

Lester and Sally Entin Faculty of the Humanities

Department of Linguistics

THREE NPS WALK INTO A SENTENCE: CAN AGREEMENT AND RESUMPTION IMPROVE CENTER EMBEDDING SENTENCES?

MA thesis submitted by

Hila Davidovitch

Prepared under the guidance of

Prof. Aya Meltzer-Asscher

October 2021

TABLE OF CONTENTS

ABSTRACT

Center Embedding sentences, such as 'The salmon that the man that the dog bit smoked tasted good', which contain two nested object-relative clauses, are notoriously difficult to process (Chomsky & Miller, 1963; Baltin & Collins, 2008). Two main explanations have been offered for this difficulty. Gibson (1998) argues that it stems from maintenance and integration costs: prohibitively high integration costs at the second verb, exceed the working memory capacity of most comprehenders, who subsequently fail in assigning the fillers to their corresponding verbs. In contrast, Lewis & Vasishth (2005) claim that the difficulty in these structures arises at retrieval: in the absence of sufficient cues, retrieval of the filler at the verb site fails due to the similarity between the three NPs, leading to interference.

This study focuses on Hebrew center embedding sentences and examines whether their perceived comprehensibility and de facto comprehension can benefit from the presence of: (i) agreement features differentially marking the three NPs and identifying every verb's subject, and (ii) resumptive pronouns (grammatical and rather freely used in Hebrew), which can aid retrieval by allowing more processing time, and/or by exhibiting the fillers' agreement features, thus unambiguously identifying the verb's object.

Experiment 1 (160 participants) addressed this question using a comprehensibility rating task. It included four conditions crossing the factors DISTINCT AGREEMENT (agreement features on the three subject NPs are all identical vs. all different) and RESUMPTION (verb objects are either gaps or resumptive pronouns). Participants read the sentences at their own pace and rated their comprehensibility on a 1-7 scale. Results revealed that neither DISTINCT AGREEMENT nor RESUMPTION significantly affected comprehensibility. There was a significant interaction between the two factors (p=.03), signaling an advantage of distinct agreement only in the absence of resumption.

Experiment 2 (192 participants) used end-of-sentence comprehension questions. Experimental sentences were of the same four conditions as in Experiment 1. The comprehension questions manipulated VERB QUESTION, targeting either the first (most embedded) or second verbs' objects. Sentences were presented word by word at a rate of 400ms per word + 200ms inter-stimulus interval. Results showed that DISTINCT AGREEMENT significantly improved comprehension (p=.004), while RESUMPTION did not. The interaction between the two factors was non-significant, meaning the cancelling-out effect resumptive pronouns had on the advantage of distinct agreement was not observed, contrary to Experiment 1. Results also revealed an effect of VERB QUESTION

(p=.001), such that the most embedded verb (and the resolution of its object dependency), presented the most difficulty. The interaction between VERB QUESTION and DISTINCT AGREEMENT was significant (p=.001), showing that while resolution of the dependency at the most embedded verb, and hence its comprehension, was not aided by distinct agreement, distinct agreement did aid the comprehension of the second verb.

The results of Experiment 2 suggest that center embedding sentences are comprehensible to some extent, especially given aid by distinct agreement. In contrast, resumption, though potentially identifying each verb's object unambiguously, did not help comprehension. These results suggest either that resumptive pronouns are not used by comprehenders for retrieval, or that interference had arisen already during the encoding of the three similar NPs (Gordon, Hendrick & Johnson, 2004; Villata, Tabor & Franck, 2018), rendering the fillers not sufficiently distinct for successful retrieval at the verb. Resumption was not only unhelpful, but it also cancelled out the advantage offered by distinct agreement in Experiment 1. This finding can be explained similarly to the 'missing V2' effect, the observation that center embedding is better accepted when only two of the three verbs appear (Frazier, 1985; Gibson & Thomas, 1999). Gibson & Thomas suggest that in such cases one of the dependencies is compromised, thus concealing the processing difficulty. Adopting this idea, it can be assumed that resumption blocks the option to neglect one of the dependencies, leading to decreased ratings.

1 INTRODUCTION

Research on sentence processing aims to characterize the mechanisms which underlie language comprehension in real time. Successful processing and comprehension rely on encoding and maintaining linguistic representations in working memory, and later on retrieving and integrating them into the incrementally built representation of the entire sentence. Much of the work in the field has focused on the operation of these processes in structures where non-adjacent elements need to be integrated, in particular filler-gap dependencies. In these dependencies, a phrase, the filler, appears clause-initially, but is interpreted in a downstream thematic position, the gap.

One way to uncover how these processes operate is to examine what causes them to fail. When encoding or maintaining a filler's features, or when retrieving these features at the integration site fail, processing difficulty might arise and comprehension might be impaired. One such notorious case is that of center embedding sentences.

1.1 CENTER EMBEDDING

Center embedding sentences, such as that in (1), consist of (at least) two nested object-relative clauses. These sentences contain three successive NPs followed by three verbs, meaning there are maximally five dependencies to resolve, as illustrated below:

(1) a. The salmon$_k$ [that the man$_j$ [that the dog$_n$ bit $\_j$] smoked $\_k$] tasted good.
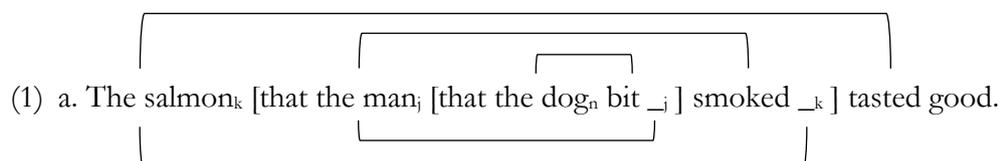
Three subject-verb dependencies need to be resolved at the verb site: 'bit' needs to be linked to 'the dog'; 'smoked' needs to be linked to 'the man'; 'tasted' needs to be linked to 'the salmon'. The two remaining dependencies that need to be resolved arise at the (in this example, gapped) verb object position: the object of 'bit' needs to be linked to its antecedent 'the man' and the object of 'smoked' needs to be linked to its antecedent 'the salmon'.

When confronted with such sentences, speakers very consistently report comprehension failure and even perception of ungrammaticality. These sentences are thus generally agreed to be difficult to process and comprehend (Chomsky & Miller, 1963; Baltin & Collins, 2008).

The accepted view, following Chomsky's distinction between competence and performance, is that there are no constraints (quantitative or combinatory) on embedding (Chomsky, 1956, 1965). In principle, the sentences in (2) become increasingly complex with the addition of each embedded

clause, so that (2c) should be somewhat harder than (2b), similarly to (2b) being somewhat harder than (2a). However (2c) is so complex that it is essentially un-processable (Gibson, 2000). Center embedding is thus a hallmark phenomenon demonstrating the limitations imposed by performance.

(2) a. The reporter disliked the editor.

   b. The reporter s'[who the senator attacked] disliked the editor.

   c. # The reporter s'[who the senator s'[who John met] attacked] disliked the editor.

1.1.1 Experimental research on center embedding

Experimental studies on center embedding have mainly gathered complexity, comprehensibility or acceptability ratings. An early study by Hamilton and Deese (1971) measured comprehensibility percentages in 3 sentence configurations: center embedding forms (e.g. "The choir that the organist that the congregation complimented directed sang new hymns"), right branching forms (e.g. "The congregation complimented the organist that directed the choir that sang new hymns") and "mixed" forms (e.g. "The choir that the organist directed that the congregation complimented sang new hymns"). Participants listened to the sentences and categorized them as either comprehensible or not. The center embedding sentences received the lowest percentages overall.

Few studies have tested de facto comprehension. Stolz (1967) used a comprehension task in which participants were read sentences and asked to break each one down into its component clauses, in the form of a list of simple sentences. Successful comprehension was assumed in cases where all verbs appeared, each with its correct subject and object. Results showed that half of the center embedding sentences were not correctly analyzed and fully comprehended. Schlezinger (1975) tested comprehension of center embedding sentences in Hebrew. Participants read complex sentences and were presented with two wh-questions following each one. Each question targeted both the subject and object of one of the embedded verbs. Participants had to fill in blanks surrounding the relevant verb. Every sentence was scored on a scale of 0 to 4, reflecting how many nouns were correctly filled in. Center embedding sentences received the lowest scores, with a mean of 1.71.

Interestingly, some findings suggest not all center embedding sentences are equally difficult. For example, center embedding sentences containing a pronoun or proper noun as their third NP, as shown in (3) and (4) respectively, are judged as less complex than those containing three full NPs (Warren & Gibson, 1999):

(3) A book [that some Italian [that **I** have never heard of] wrote] will be published soon by MIT press.

(4) The reporter [who the senator [who **John** met] attacked] disliked the editor.

In addition, there is evidence that this difficulty is less extreme in other languages. Vasishth, Suckow, Lewis & Kern (2010), showed better comprehension of center embedding sentences in German compared to English: German speakers answered comprehension question with 65% average accuracy rates, while English speakers answered with 54% accuracy rates.

Considering theoretical frameworks have since been updated and experimental methodologies have developed and become more reliable, it seems worth re-examining the extent to which center embedding sentences are in fact incomprehensible. Before turning to the current study, a brief review of prevalent accounts of findings on center embedding is presented.

1.1.2   Accounts of center embedding

The first attempt to capture the complexity of center embedding sentences and explain their standout difficulty is attributed to Yngve (1960). He assumed a limited capacity working memory and surmised that the difficulty of center embedding sentences stems from failure to retain a high number of unresolved syntactic dependencies. Consider sentence (5):

(5) The boy [who the teacher [who the neighbor saw] met] fell.

After processing 'neighbor', there are five unresolved dependencies: 'The boy' awaits a verb for which it will be the subject; the first 'who' awaits a verb for which it will be the object; 'the teacher' awaits a verb for which it will be the subject; the second 'who' awaits a verb for which it will be the object; and 'the neighbor' awaits a verb for which it will be the subject.

In the same vein as Yngve's analysis, Lewis (1996) claims the difficulty rises from maintaining open dependencies of the same type. Specifically, at the point of 'neighbor', there are three nominative-marked NPs ('the boy', 'the teacher', 'the neighbor') awaiting a verb's subject position to fill.

More recently, two main accounts have been proposed to depict the breakdown in center embedding sentences, each capturing the difficulty based on a different component of processing.

Gibson (1998) proposed the Dependency Locality Theory (DLT) to describe the use of computational resources in sentence processing and comprehension. He claims that
resources are required for two aspects of constructing an interpretation for a sentence:
 (i) Maintenance of an already built structure, which includes keeping track of unresolved dependencies.

(ii) Integration of each word into the built structure, which has two components: (a) structural integration, which is a predicative process connecting a word into the already built structure, and (b) discourse integration, during which discourse referents are constructed or accessed.

Gibson attributes the difficulty in center embedding sentences to both maintenance and integration costs. As far as maintenance, he suggests that the memory load associated with keeping track of the three fillers (nouns) and the predictions associated with them exceeds the working memory capacity of most comprehenders. Subsequently, the parser fails in assigning the fillers to their corresponding verbs and the resolution of the dependencies fails. As for integration, costs are calculated using energy units (EUs). Discourse integration requires constructing a new discourse referent for every verb and (lexical) noun in the sentence, consuming one EU per referent. Structural integration complexity depends on the distance between the two words being integrated. It rises as the distance grows and more EUs are consumed per each new discourse referent in the intervening region, leading to incremented costs. According to Gibson, a sentence's perceived complexity is largely determined by the local maximal integration cost. For example, in sentence (2c), presented here again as (6), integration costs reach a maximum at the second verb:

(6) The reporter $_s$'[who the senator $_s$'[who John met] attacked] disliked the editor.

The processing of 'attacked' consumes 7 EUs: 1 to build the verb's discourse referent; 2 to link it to its subject, 'the senator', across two intervening referents: 'met' and 'John'; and 4 to link the object position gap to its antecedent, 'the reporter', across 4 intervening referents: 'attacked', 'met', 'John' and 'the senator'.

The DLT accounts for the previously mentioned finding regarding reduced complexity of center embedding sentences containing pronouns, as in (3). Gibson assumes that 1st and 2nd person pronouns do not require constructing a discourse referent, thus reducing the maximal integration cost which translates into lower complexity ratings.

Another prominent account was suggested by Lewis & Vasishth (2005) in the cue-based retrieval framework. In this model, linguistic elements are encoded and stored as bundles of feature-value pairs. During processing, incoming words which need to be integrated into the existing structure trigger a search for specific previously encountered constituents. This search is guided by feature values that the current dependency requires. For example, a verb will initiate a memory search for a constituent which can complete its required subject-verb dependency. When the searched features are matched to previously provided cues, successful retrieval occurs. Following Lewis (2000), the authors embrace the notion that there is no serial order representation in sentence

processing, and parsing is rather based solely on cue-based associative retrievals. For example, in the sentence (7), 'the reviewers' is encoded as a PLURAL, ANIMATE, NOMINATIVE NP:

(7) The musician [who the reviewers praise so highly] will likely win the Grammy.

Upon arriving at the verb 'praise', its subject (an animate, nominative-marked NP) needs to be retrieved. The verb also bears plural agreement, generating a PLURAL retrieval cue. 'The reviewers' perfectly matches the verb's retrieval cues and is thus very likely to be correctly retrieved as its subject.

Lewis & Vasishth (2005) suggest that the difficulty in center embedding sentences arises at the retrieval portion of processing. The basic problem with center embedding sentences is that they contain multiple attachment points that require distinguishing candidate constituents primarily or exclusively based on their relative serial order. Specifically, there are two active fillers and two predicted embedded clauses that must be properly distinguished by serial order to make the correct attachments at the verbs. However, since serial order per se is not represented, and retrieval cues at the verbs are insufficient to distinguish between the three candidates, similarity-based interference arises during retrieval, making these sentences likely to be mis-parsed. For example, in sentence (1b), presented here again as (8), the verb 'liked' only cues that it needs an ANIMATE NP subject and an NP object:

(8) The boy [who the neighbor [who the guest liked] saw] fell

At this point there are three NPs maintained in working memory, all (singular) animates. The retrieval cues provided by the verb are insufficient to distinguish between the three NPs and similarity-based interference leads to failure to identify and retrieve the correct arguments of the verb.

1.2 RESUMPTIVE PRONOUNS

Resumptive pronouns (RPs) are overt elements which appear at the tail of a wh-dependency, where a gap would otherwise appear. They appear rather freely in languages termed "grammaticized resumption" languages (Sells, 1984), e.g. Hebrew. In other, "intrusive resumption" languages, e.g. English, their occurrence is more restricted and they are generally judged as ungrammatical.

A prevalent observation in the literature is that the acceptability of resumptive pronouns increases in environments that are considered harder to process, such as islands and deeply embedded positions (Ross, 1967; Ariel, 1999). These findings have led to the claim that resumptive pronouns may serve a facilitatory processing function in such complex structures, perhaps in aiding the retrieval of a less accessible filler. The evidence for such a processing advantage in intrusive

resumptive languages is mixed, with some studies showing that resumptive pronouns can serve to repair island violations (Han, Elouazizi, Galeano, Görgülü, Hedberg, Hinnell, Kim, Kyeong-min & Kirby, 2012), and others showing that resumptive pronouns did not raise island sentences' acceptability compared to gaps (Alexopoulou & Keller, 2007). Studies on resumption in Hebrew reveal that in relatively simple to process structures resumptive pronouns elicit lower acceptability ratings (Meltzer-Asscher, Fadlon, Goldstein & Holan, 2015). They were, however, found to increase grammaticality of islands (Farby, Danon, Walters & Ben-Shachar, 2010; Keshev, 2016; Keshev & Meltzer-Asscher, 2017). This raises the possibility that in grammaticized resumption languages, resumptive pronouns may have some processing function in complex structures after all.

In the case of Hebrew center embedding sentences, resumptive pronouns could aid in resolving the dependencies, namely in retrieving the object of the verb, in two ways:

(i) Wagers & Phillips (2014) argued that in long dependencies, additional time is needed for retrieval of the filler's semantic information. Resumptive pronouns, being pronouns, are about 200-300 ms long. As successful retrieval is estimated to take 85 ms (McElree, Foraker & Dyer, 2003), resumptive pronouns could allow the comprehender the needed extra time to retrieve and integrate all the relevant information associated with the filler, before new lexical information arrives and requires processing. In sentences with no resumption, a verb would be followed by another verb, with its processing costs, higher than those of a pronoun, and therefore no extra processing time will be afforded, leading to increased difficulty.

Resumptive pronouns bear the gender and number agreement features of their fillers (McCloskey, 2006). They can therefore serve as retrieval cues, by unambiguously discriminating between potential fillers, providing they are marked with distinct agreement features. This should minimize interference and enable accurate retrieval.

2   RESEARCH QUESTIONS

As presented in the introduction, previous studies on center embedding sentences in English revealed that some factors (e.g. the use of a pronoun instead of lexical noun) can contribute to their acceptability and reduce their perceived complexity. Hebrew has unique qualities which makes it a prime candidate to further investigate such ameliorating factors.

In most languages, verbs must agree with their subjects in number, person, gender and sometimes other features in order to establish a grammatical formation (Mallinson & Blake, 1981). Hebrew, unlike English, which was the focus of most research on center embedding, shows overt marking

of gender (masculine and feminine), number (singular, plural and dual) and person on nouns and predicates, and mandates subject-predicate agreement. As agreement indicates the link between a subject and its predicate, it is essential for sentence comprehension. Overtly marking agreement features explicates the link, which provides for a facilitatory effect (Acuña-Fariña, 2009). In addition, as explained above, Hebrew, unlike English, has grammaticized resumptive pronouns, which likewise may facilitate processing of dependencies.

Given this, the current study aims to explore the following questions:

1. Do distinct agreement markings and resumptive pronouns indeed serve as effective retrieval cues? If so, under what conditions? Distinct agreement marking should in principle unambiguously identify each verb's subject, while resumptive pronouns should unambiguously identify each verb's object, thus potentially aiding correct retrieval and integration of all the arguments in the sentence.

2. Is there an interaction between distinct agreement and resumptive pronouns in the processing and comprehension of center embedding sentences?

3. Are center embedding sentences truly beyond comprehension or can they be somewhat comprehensible providing sufficient aids? Will these sentences be perceived as more easily comprehensible in the presence of distinct agreement marking and resumption pronouns? Will they in fact be somewhat comprehended?

3   EXPERIMENT 1

This experiment used a comprehensibility rating task. It included 4 conditions manipulating:

(i)   AGREEMENT: all three NPs (and consequentially the three verbs) were either all identical in terms of the combination of gender and number features (SAME AGREEMENT), or all different (DISTINCT AGREEMENT).

(ii)   RESUMPTION: verb objects were either gaps (NO RP) or resumptive pronouns (RP).

The contrast between (DISTINCT AGREEMENT, NO RP) and (SAME AGREEMENT, NO RP) tests whether distinct agreement marking increases the perceived comprehensibility of the sentences. When the three NPs bare distinct agreement features, agreement on the verb should unambiguously identify each subject and aid in the retrieval of the subject NPs.

The contrast between (SAME AGREEMENT, RP) and (SAME AGREEMENT, NO RP) tests whether the mere extra time afforded by resumptive pronouns increases the perceived comprehensibility of the sentences.

However, resumptive pronouns could also be found to increase comprehensibility only in the presence of distinct agreement, where their features are informative regarding the identity of the filler. The contrast between (DISTINCT AGREEMENT, RP) and (SAME AGREEMENT, NO RP) tests this. When the three NPs bare distinct agreement features and objects are resumptive pronouns, agreement on the verb should unambiguously identify each verb's subject, and resumptive pronouns should unambiguously identify each verb's object. This should aid correct retrieval of all the NPs in the sentence.

3.1 METHOD

3.1.1 PARTICIPANTS

160 participants volunteered to take part in the experiment. All were native Hebrew speakers, between the ages of 18 to 40 (average=26).

3.1.2 MATERIALS

Eight sentence sets were composed, with four conditions each, as exemplified in Table 1 below.

SAME AGREEMENT	NO RP	ha-balon.**SG-M** še-ha-leycan.**SG-M** še-ha-yeled.**SG-M** icben.**SG-M** __ nipeax.**SG-M** __ hitpocec.**SG-M**
	RP	ha-balon.**SG-M** še-ha-leycan.**SG-M** še-ha-yeled.**SG-M** icben.**SG-M oto** nipeax.**SG-M oto** hitpocec.**SG-M**
Hebrew		הבלון שהליצן שהילד עצבן (אותו) ניפח (אותו) התפוצץ
'The balloon that the clown that the child annoyed (him) inflated (him) popped'		
DISTINCT AGREEMENT	NO RP	ha-balon.**SG-M** še-ha-leycanit.**SG-F** še-ha-yeladim.**PL-M** icbenu.**PL-M** nipxa.**SG-F** hitpocec.**SG-M**
	RP	ha-balon.**SG-M** še-ha-leycanit.**SG-F** še-ha-yeladim.**PL-M** icbenu.**PL-M ota** nipxa.**SG-F oto** hitpocec.**SG-M**
Hebrew		הבלון שהליצנית שהילדים עצבנו (אותה) ניפחה (אותו) התפוצץ
'The balloon that the clown(f) that the children annoyed (her) inflated (him) popped'		

Table 1: Example set for experimental conditions

Each participant saw one sentence from each set, in a Latin Square design. Thus, each participant saw only two sentences from each condition. This was done in order to avoid an effect of adaptation (Fine, Jaeger, Farmer, & Qian, 2013) to the highly marked center embedding structure and to prevent participants from (unconsciously) developing a processing strategy

Twenty-four filler sentences of three types were used (eight sentences per type), as exemplified in Table 2 below. Each type of filler had a distinct-agreement features variation and a same-agreement features variation (four sentences per variation), in order to distract from the agreement-related experimental manipulation. Sixteen of the sentences contained resumptive pronouns in order to distract from the resumption-related experimental manipulation. All fillers contained two relative clauses, both object and subject ones, in order to bring the fillers as close to the complexity level of the experimental sentences as possible.

Type 1	ha-šir še-hiršamti et ha-amargan še-šama oto hitnagen ba-reka
	'The song that I impressed the manager who heard it played in the background'
Type 2	ha-kafe še-šatiti im ha-baxur še-ha-kupa'i hirgiz nišpax
	'The coffee that I drank with the guy that the cahier angered spilled'
Type 3	ha-ca'acu'a še-ha-pa'ot še-ibed oto baxa hitgalgel el ha-sixim
	'The toy that the infant who lost it cried rolled into the bushes'

Table 2: Example for different types of filler sentences

3.1.3   PROCEDURE

The experiment was conducted online via Google Forms. Participants were presented with instructions (provided in appendix B) to read each sentence and rate its level of comprehensibility on a scale of 1-7, with 1 being "completely incomprehensible" and 7 being "easily comprehensible". 2 examples were given to familiarize participants with the scale. All sentences (experimental and filler) were presented in one page, in randomized order, with each experimental sentence being followed by three filler sentences.

3.2   RESULTS

Raw comprehensibility ratings for each condition are presented in Table 3 and Figure 1 below.

Condition	Raw average rating
SAME AGREEMENT, NO RP	3.4
SAME AGREEMENT, RP	3.3
DISTINCT AGREEMENT, NO RP	3.9
DISTINCT AGREEMENT, RP	3.5

**Table 3:** Raw average comprehensibility ratings
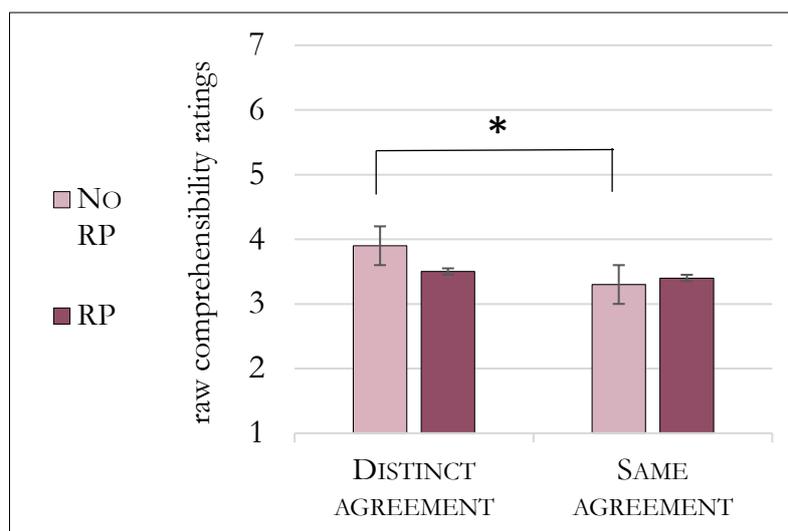
Figure 1: Mean comprehensibility ratings (error bars represent +/-1 standard error of the mean)

Data were analyzed with mixed effects models. Analyses were conducted using the lmerTest package (Kuznetsova, Brockhoff & Christensen, 2014) in the R software environment (R Development Core Team, 2011).

No main effect was found for either AGREEMENT or RESUMPTION. A significant interaction between AGREEMENT and RESUMPTION was found, such that DISTINCT agreement improved comprehensibility only in the absence of resumption (p=.027).

3.3 DISCUSSION

Sentences with distinct agreement features on the NPs received higher ratings than those with the same agreement features, but only in the absence of resumption. Despite the lack of a main effect for agreement, this interaction indicates that distinct agreement features may indeed somewhat help overt identification of verbs' subjects. This aid improves comprehensibility, but resumption detracts from this effect.

The fact that no main effect was found for resumption could suggest that resumptive pronouns do not provide sufficient cues for retrieval. This is perhaps because the filler's features are not maintained stably enough to begin with (Wagers, Lau & Phillips, 2009), so even an unambiguous cue is not helpful. Furthermore, as previously mentioned, even though Hebrew is a grammatical resumption language, it appears that in relatively simple sentences resumptive pronouns reduce acceptability ratings. They do, however, increase the acceptability of sentences with islands. It is possible that resumptive pronouns can only serve to repair ungrammatical structures and cannot aid in ameliorating difficulties arising from the processing of grammatical sentences, however complex. This may be due to the fact that the resumptive pronouns occur "too late" in the

sentence. Retrieval is already attempted at the verb, if interference arises at this stage and retrieval fails, the resumptive pronoun, which is encountered after retrieval failure, cannot salvage the process.

It is also worth noting that all experimental sentences contained two animate NPs and an inanimate NP. Returning to Lewis & Vasishth's (2005) proposal that the difficulty in center embedding sentences stems from failed retrieval due to similarity-based interference, it seems that if this was indeed the source of difficulty, the experimental sentences would perhaps be perceived as more comprehensible then what was actually found, at the very least in the presence of distinct agreement. This is because one of the NPs is distinguished from the other two by (non-)animacy, and perhaps other semantic features, reducing interference and aiding retrieval at the verb. For example, in the set provided in Table 1, given the verb 'inflated', only the inanimate NP 'the balloon' is a semantically appropriate object for the verb. However, comprehensibility ratings were very low, suggesting that similarity-based interference is not the factor leading to failed retrieval, and perhaps that failed retrieval is not the source of the difficulty at all.

4 EXPERIMENT 2

After investigating perceived comprehensibility in the previous experiment, the current experiment set out to investigate whether center embedding sentences do involve such a severe processing breakdown that comprehension completely fails.

The current experiment used end-of-sentence comprehension questions. Like in Experiment 1, it included a manipulation of the distinctiveness of the three NPs' agreement features and the presence of resumptive pronouns. Comprehension questions were asked about two of the verbs in the sentence (this will be elaborated upon in the material section), thus also manipulating VERB QUESTION, yielding eight conditions overall.

4.1 METHOD

4.1.1 PARTICIPANTS

192 subjects participated in the experiment. Some received 15 ILS for their participation and others received course credit in Tel Aviv University's Linguistics department. All were native Hebrew speakers, between the ages of 18 to 35 (average=25).

Similarly to Experiment 1, eight sentence sets were composed, with four conditions each, as exemplified in Table 4 below. Experimental sentences were reformulated from the previous experiment to include animate nouns only, so that verbs could not be matched with their subjects or objects by using only semantic cues. The reason to eliminate such biases is that they can promote comprehension of the sentence without thoroughly processing it, and perhaps without even fully constructing its syntactic structure. Also similarly to Experiment 1, each participant saw only two sentences from each condition, in order to prevent adaptation.

SAME AGREEMENT	NO RP	ha-yeled.**SG-M** še-ha-šaxen.**SG-M** še-ha-oreax.**SG-M** hivhil.**SG-M** __ xibev.**SG-M** __ nafal.**SG-M**
	RP	ha-yeled.**SG-M** še-ha-šaxen.**SG-M** še-ha-oreax.**SG-M** hivhil.**SG-M oto** xibev.**SG-M oto** nafal.**SG-M**
Hebrew		הילד שהשכן שהאורח הבהיל (אותו) חיבב (אותו) נפל
'The child that the neighbor that the guest frightened (him) liked (him) fell'		
DISTINCT AGREEMENT	NO RP	ha-yeled.**SG-M** še-ha-šxenim.**PL-M** še-ha-oraxat.**SG-F** hivhila.**SG-F** xibevu.**PL-M** nafal.**SG-M**
	RP	ha-yeled.**SG-M** še-ha-šxenim.**PL-M** še-ha-oraxat.**SG-F** hivhila. **SG-F** **otam** xibevu.**PL-M** **oto** nafal.**SG-M**
Hebrew		הילד שהשכנים שהאורחת הבהילה (אותם) חיבבו (אותו) נפל
'The child(m) that the neighbors that the guest(f) frightened (them) liked (him) fell'		

Table 4: Example set for experimental conditions

The comprehension questions targeted the object of either the most embedded verb (V1, *hivhil/a*, 'frightened' in the set in Table 4) or the second verb (V2, *xibevu*, 'liked' in the set in Table 4), as exemplified in Table 5 below. Each participant answered one V1 question and one V2 question per condition. Questions following DISTINCT AGREEMENT sentences contained the same agreement markings as in the sentence, to make them natural in the context. Two possible answers were provided, consisting of the two NPs which appeared in the sentence but were not the subject of the verb in the question. The NPs were presented one above the other for participants to choose between. Half of the correct answers appeared as the top option and the other half as the bottom to avoid a choice preference or bias. No feedback was given so that participants would not use it to develop a strategy.

	SAME AGREEMENT	et mi ha-oreax hivhil?	
		Whom (did) the-guest.**SG-M** frighten.**SG-M**	
V1	DISTINCT AGREEMENT	et mi ha-oraxat hivhila?	
		Whom (did) the-guest.**SG-F** frighten.**SG-F**	
	Answers	ha-yeled (the child)	**ha-šaxen/šxenim** (the neighbor/s)
	SAME AGREEMENT	et mi ha-šaxen xibev?	
		Whom (did) the-neighbor.**SG-M** like.**SG-M**	
V2	DISTINCT AGREEMENT	et mi ha-šxenim xibevu?	
		Whom (did) the-neighbors.**PL-M** like.**PL-M**	
	Answers	**ha-yeled** (the child)	ha-oreax/oraxat (the guest/s)

Table 5: Example comprehension questions and possible answers for the set exemplified in Table 4 above (correct in bold)

Sentences and questions were divided to lists in a Latin Square design, so that each participant only saw one sentence per set, with one comprehension question. This resulted in 8 experimental lists.

Similarly to Experiment 1, twenty-four filler sentences of three types were used, each with a distinct-agreement features variation and a same-agreement features variation. All sentences contained both an object relative clause and a subject relative clause. Unlike Experiment 1, all fillers contained resumptive pronouns. Comprehension question, about verbs' objects and subjects, were asked following each of the filler sentences so as to not distinguish them from the experimental sentences.

Type 1	ha-ciur še-he'eraxti et ha-aman še-ciyer oto nimkar 'The painting that I admired the artist who painted it was sold'
Type 2	ha-bakbuk še-ha-kelev še-la'as oto navax neheras 'The bottle that the dog that chewed it barked was destroyed'
Type 3	ha-poše'a še-ha-balaš še-tafas oto hudax hištaxrer 'The criminal that the detective who caught him was dismissed was released'

Table 6: Example for all types of filler sentences

4.1.3 PROCEDURE

The experiment was conducted online via the now defunct Ibex Farm platform[1]. Participants were presented with instructions (provided in appendix D) to read the sentences and then answer a question that would appear after the sentence had ended. Participants were instructed to choose the most suitable answer out of two possible ones. Sentences were presented using rapid serial visual presentation, with each word presented in the center of the screen for 400ms before disappearing, and the following word appearing after a 200ms interval. Four practice trials were conducted to help participants accommodate to the presentation method. After the sentence ended, the question and two possible answers appeared on the screen. Participant had an unlimited amount of time to read and answer the question while it and the possible answers remained on the screen.

4.2 RESULTS

Raw average accuracy rates for each condition are presented in Table 7 and Figure 2 below.

Condition	Raw average rates
SAME AGREEMENT, NO RP	0.53
SAME AGREEMENT, RP	0.56
DISTINCT AGREEMENT, NO RP	0.60
DISTINCT AGREEMENT, RP	0.63

Table 7: Raw average accuracy rates


[1] https://spellout.net/ibexfarm/
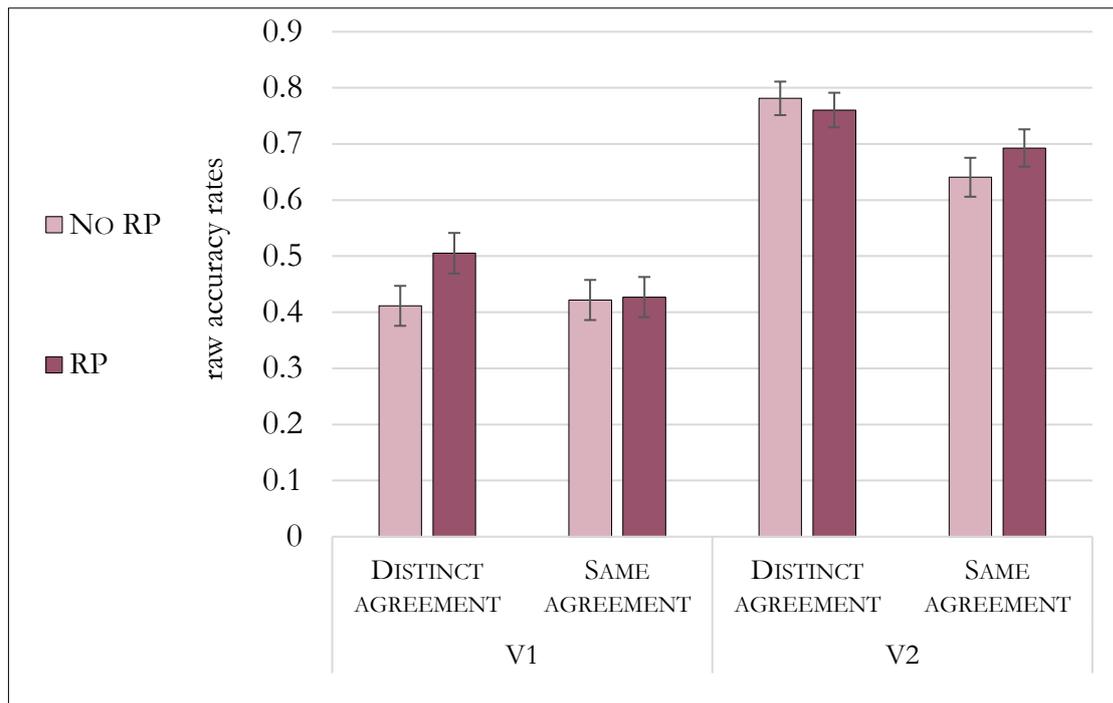
Figure 1: Mean accuracy rates (error bars represent +/-1 standard error of the mean)

Data were analyzed with mixed effects models. Analyses were conducted using the lmerTest package in the R software environment.

No main effect or interactions were found for RESUMPTION. A main effect was found for AGREEMENT (p=.0036), such that DISTINCT agreement improved overall accuracy rates. A main effect for VERB QUESTION was also found (p=.001), such that accuracy rates were higher for V2 questions. Finally, a significant interaction was found between AGREEMENT and VERB QUESTION (p=.027), such that DISTINCT agreement aided comprehension of V1, but not of V2.

4.3   DISCUSSION

The results of this experiment establish that distinct agreement provides an advantage for comprehension: on average, same agreement sentences had 55% accuracy rates, while distinct agreement sentences had 62% accuracy rates, a significant 7% difference. Since comprehension questions targeted the embedded verbs' objects, this finding indicates that distinct agreement aids not only in correctly identifying the verbs' subjects but also their objects.

The finding from Experiment 1 concerning resumption and its and lack of processing facilitation has replicated (p=.28). However, the cancelling-out effect it had on the advantage of distinct agreement had not replicated. In the current experiment there was no interaction between agreement and resumption, with resumption taking away the advantage of agreement.

Furthermore, the results suggest that overall, V1, the most embedded verb, posed the biggest obstacle for comprehension, while V2 was less problematic in comparison (p<.0001). Firstly, on average, V1 question had 44% (below chance level) accuracy rates, while V2 question had 72% accuracy rates (well above chance level), a rather staggering 28% difference. This unequivocally means it was easier for participants to answer questions about the object of V2 than about the object of V1. Secondly, the finding that distinct agreement increased accuracy for V2 questions but not for V1 questions further supports the impression that the difficulty is more strongly associated with the most embedded verb.

5 GENERAL DISCUSSION

In the heart of the line of research on center embedding lies the notion that this structure is so complex that it "often leads to a breakdown" (Baltin & Collins, 2008). The aim of this study was to uncover whether this is truly the case, and what, if anything, can prevent this breakdown. Results of two experiments testing perceived comprehensibility and comprehension paint a more nuanced picture. These findings join those mentioned in the intro and together they suggest that the processing and comprehension of center embedding sentence does not necessarily involve a complete and unavoidable breakdown.

5.1 (Non-)Effects of resumption

A consistent finding in both experiments is the lack of a contributing effect to resumption: the occurrence of resumptive pronouns in object position did not increase comprehensibility ratings or comprehension accuracy rates. This could suggest that resumptive pronouns, and the extra processing time they afford, do not provide sufficient retrieval aid. Their presence and the extra processing time stemming from it could even be considered a hinderance, given the cancelling-out effect resumption had on distinct agreement in experiment 1.

This hindering influence of resumptive pronouns can be accounted for similarly to the 'missing V2' effect, exemplified in sentence (9) below. This effect is a well-known linguistic illusion, namely the observation that center embedding sentences are judged as more acceptable when one of verbs - specifically the second - does not appear, despite their resulting ungrammaticality (Frazier, 1985; Gibson & Thomas, 1999).

(9) * The patient [who the nurse [who the clinic had hired] met Jack.

Gibson & Thomas (1999) suggest that in missing V2 configurations, one of the dependencies is compromised, due to the prediction for that verb being forgotten under a heavy working memory

load. This ameliorating effect conceals the processing difficulty associated with the structure. Adopting this idea, it can be assumed that resumption blocks the option to neglect one of the dependencies, leading to decreased comprehensibility ratings.

The rating task itself could have also contributed to the manifestation (or lack thereof) of this cancelling-out effect. Providing judgments, i.e. making an offline, conscious decision about some quality of a sentences, can be affected by interfering factors, in this case - the length added by the presence of a resumptive pronoun. Such considerations do not come into play during online comprehension and are therefore less likely to affect performance on comprehension tasks. This cancelling-out effect could therefore be task specific, explaining why it did not replicate in experiment 2.

That being said, in both experiments resumptive pronouns did not aid retrieval when all NPs, and therefore they themselves, were marked with same agreement features. It is perhaps not surprising, as it could be argued that in the absence of distinct agreement marking, resumptive pronouns cannot actually distinguish between competing filler candidates, as they lack the prime means for such disambiguation, i.e. the overt marking. This will be further explored when discussing future research goals.

The results from the experiments weaken Lewis & Vasishth's (2005) proposal that center embedding sentences are difficult due to insufficient discriminating cues at retrieval. Since in the distinct agreement condition resumptive pronouns unambiguously marked the correct filler, they should have had at least some contribution to participants' performance, contrary to fact.

An alternative tentative explanation for why resumptive pronouns, though potentially identifying each verb's object unambiguously, did not help comprehension, is that interference had arisen already during the encoding of the three similar NPs (Gordon, Hendrick & Johnson, 2004; Villata, Tabor & Franck, 2018), rendering the fillers not sufficiently distinct for successful retrieval at the verb.

5.2 Verb question effect

Perhaps the most surprising finding, from Experiment 2, is that answering questions about the object of V1 was significantly more difficult than answering questions about the object of V2. In fact, for V2-object questions, accuracy rates were above chance levels, at 72% accuracy, a rather high rate considering the complexity of center embedding sentences.

This finding could be considered to stand in contrast to the observed missing V2 effect. The fact that it is specifically the second verb whose omission enhances acceptability and illudes

grammaticality would seem to suggest that it is the source of the processing difficulty associated with the structure, in contrast to Experiment 2's findings, showing rather successful comprehension of the thematic role associated with V2. It is worth noting, however, that it is not clear whether the missing V2 effect is consistent cross-linguistically. For example, there is some evidence to support that German comprehenders do not experience this illusion (Vasishth et al., 2010). Based on the current study's findings, Hebrew could present a similar case to German.

The finding regarding the relative ease of answering V2 questions leads to the impression that the difficulty in center embedding sentences is more strongly associated with the most embedded verb, V1. Gibson's (2000) DLT model, however, associates the difficulty to incremented integration costs localized to the second verb. The conflict between these observations could potentially also be resolved by turning to examine the tasks under which each finding was observed. Gibson tested complexity ratings and argued that higher ratings correlate to the maximal local integration cost. It could be that there is indeed a considerable processing difficulty of the second verb, which manifests in ratings. This difficulty could dissipate as processing continues and the time comes to make a comprehension decision. It could also be that said difficulty is irrelevant for answering comprehension questions.

Interesting observations arise when examining the effects of agreement and resumption on the two different verb questions separately. Table 8 below provides the raw average accuracy rates per verb.

Condition	Accuracy rates	Verb accuracy
SAME AGREEMENT, NO RP, V1	0.42	
SAME AGREEMENT, RP, V1	0.43	0.44
DISTINCT AGREEMENT, NO RP, V1	0.41	
DISTINCT AGREEMENT, RP, V1	0.51	
SAME AGREEMENT, NO RP, V2	0.64	
SAME AGREEMENT, RP, V2	0.69	0.72
DISTINCT AGREEMENT, NO RP, V2	0.78	
DISTINCT AGREEMENT, RP, V2	0.76	

Table 8: Raw average accuracy rates per verb

The three-way interaction between AGREEMENT, RESUMPTION and VERB QUESTION was not significant. However, it can be observed that in V2 questions, in the absence of resumptive pronouns, distinct agreement has a substantial facilitatory influence (78% vs. 64%, a 14% increase), whereas in the presence of a resumptive pronoun, this facilitation decreases (76% vs. 69%, only a

7% increase due to distinct agreement). This pattern, wherein resumption cancels out the advantage of agreement, could be considered a replication of the similar finding from experiment 1. This once again leads in the direction of resumptive pronouns failing to serve as retrieval aids. However, the pattern observed for V1 questions is worth further investigation: in V1 questions with same agreement marking, resumptive pronouns seem to have no effect on participants' accuracy levels, but in the presence of distinct agreement they do appear to enhance accuracy rates. Participants' performance in these questions displays, for the first time in the current study, signs of a trend such that resumptive pronouns do assist distinct agreement in retrieval, finally serving their hypothesized purpose. However, since this finding was limited to V1 questions and is based on a small number of trials per participant, it is not very reliable. In order to replicate this finding and further explore it, I propose two experiments, described below.

6 FUTURE RESEARCH

A way to establish a facilitating effect of distinct agreement-marked resumptive pronouns on comprehensibility and comprehension, is by testing whether the occurrence of only one resumptive pronoun in a sentence may offer even more aid. It could be that the effect of resumption did not manifest in the experiments presented in this paper simply due to the fact that one resumptive pronoun would have sufficed to aid identifying and retrieving a filler, but two resumptive pronouns in one sentence are a hindrance. Furthermore, the different patterns of resumption behavior between the different verb questions may suggest that resumptive pronouns are needed or are of benefit only in certain positions, and unnecessary or hindering in others.

The experiments will use the tasks of Experiments 1 and 2 above (namely a comprehensibility ratings experiment and a comprehension experiment). Materials will be similar to those of experiment 2, but reformulated such that two sentences will contain only one resumptive pronoun each, either in the object of V1 position or in the object of V2 position. A sentence with no resumptive pronouns and a sentence with two resumptive pronouns will also be utilized, in order to allow a more reliable comparison to the experiments presented in this paper. Furthermore, providing that resumptive pronouns can maximally aid retrieval when they are distinctly marked, no manipulation of agreement will be performed and the NPs in the sentence, and hence the resumptive pronouns, will be marked with distinct agreement features. An example set is provided in Table 9 below.

NO RP	ha-yalda.**SG-F**	še-ha-šxenim.**PL-M**	še-ha-oreax.**SG-M**
	hivhil.**SG-M** hikiru.**PL-M** nafala.**SG-F**		
RP V1	ha-yalda.**SG-F**	še-ha-šxenim.**PL-M**	še-ha-oreax.**SG-M**
	hivhil.**SG-M** otam hikiru.**PL-M** nafala.**SG-F**		
RP V2	ha-yalda.**SG-F**	še-ha-šxenim.**PL-M**	še-ha-oreax.**SG-M**
	hivhil.**SG-M** hikiru.**PL-M** ota nafala.**SG-F**		
2 RPS	ha-yalda.**SG-F**	še-ha-šxenim.**PL-M**	še-ha-oreax.**SG-M**
	hivhil.**SG-M** otam hikiru.**PL-M** ota nafala.**SG-F**		
HEBREW	הילדה שהשכנים שהאורח הבהיל (אותם) הכירו (אותה) נפלה		
'The girl that the neighbors that the guest(m) frightened (them) knew (her) fell'			

Table 9: Example set for experimental conditions

A variation between two possible orders of features will be implemented: in half of the experimental sets the first NP will be marked with FEMININE SINGULAR features and the second NP will be marked with MASCULINE PLURAL features (as shown in Table 9 above); in the other half of experimental sets that order will be reversed. This will also affect the order of features presented on the resumptive pronouns. An example set for the alternative feature order is provided in Table 10 below. Variating these two orders is necessary to ensure that any (hypothetically) observed effect would have to do only with the position in which the resumptive pronoun appears and not with the agreement features themselves.

NO RP	ha-morim.**PL-M**	še-ha-talmida.**SG-F**	še-ha-menahel.**SG-M**
	xibev.**SG-M** tsiyra.**SG-F** huxme'u.**PL-M**		
RP V1	ha-morim.**PL-M**	še-ha-talmida.**SG-F**	še-ha-menahel.**SG-M**
	xibev.**SG-M** ota tsiyra.**SG-F** huxme'u.**PL-M**		
RP V2	ha-morim.**PL-M**	še-ha-talmida.**SG-F**	še-ha-menahel.**SG-M**
	xibev.**SG-M** tsiyra.**SG-F** otam huxme'u.**PL-M**		
2 RPS	ha-morim.**PL-M**	še-ha-talmida.**SG-F**	še-ha-menahel.**SG-M**
	xibev.**SG-M** ota tsiyra.**SG-F** otam huxme'u.**PL-M**		
HEBREW	המורים שהתלמידה שהמנהל חיבב (אותה) ציירה (אותם) פרשו		
'The teachers that the pupil(f) that the principle(m) liked (them) drew (her) retired'			

Table 10: Example set for alternative feature order experimental conditions

If higher accuracy rates are found in conditions with one resumptive pronoun (namely 2 and 3) compared to the condition with no resumptive pronoun (namely 1), it would indicate that a resumptive pronoun can indeed serve as a retrieval cue as to the object of a verb, when it appears in the proper position for it. Following the pattern observed in Experiment 2, where a distinctly marked resumptive pronoun numerically increased accuracy rates for V1 questions, if higher rates are found when a resumptive pronoun appears in the V1 object position (condition 2), it would indicate that resumption is indeed more beneficial for retrieval in that position. It would be interesting to see whether rates are even higher in condition 4, i.e. in the presence of two distinctly marked resumptive pronouns, or whether such resumption is excessive and ultimately hindering.

REFERNCES

Acuña-Fariña, J. C. 2009. The linguistics and psycholinguistics of agreement: A tutorial overview. *Lingua*, 119(3): 389-424.

Alexopoulou, T., & Keller, F. 2007. Locality, cyclicity, and resumption: At the interface between the grammar and the human sentence processor. *Language*, 110-160.

Ariel M. 1999. Cognitive universals and linguistic conventions: The case of resumptive pronouns. *Studies in Language*, 23: 217-69.

Baltin, M., & Collins, C. 2008. The handbook of contemporary syntactic theory, vol. 23. *John Wiley & Sons.*

Chomsky, N. 1956. On the limits of finite-state description. *MIT Research Laboratory for Electronics, Quarterly Progress Report* 41: 64–65.

Chomsky, N., & Miller, G. 1963. Introduction to the formal analysis of natural languages. In Luce, R.D., Bush, R.R., Galanter, E. (Eds.), *Handbook of Mathematical Psychology,* 2: 269–321. Wiley, New York.

Chomsky, Noam. 1965. Aspects of the Theory of Syntax. Cambridge, MA: MIT Press

Farby S., Danon G., Walters J., & Ben-Shachar M. 2010. The acceptability of resumptive pronouns in Hebrew. In *Proceedings of IATL (Israel Association for Theoretical Linguistics),* 26, ed. Y. Falk.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. 2013. Rapid expectation adaptation during syntactic comprehension. *PloS one*, 8(10).

Frazier, L. 1985. Syntactic complexity. *Natural language parsing: Psychological, computational, and theoretical perspectives*, 129-189.

Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition,* 68: 1-76.

Gibson, E., & Thomas, J. 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes*, 14(3): 225-248.

Gibson, E., 2000. The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain*, 95-126.

Gordon, P. C., Hendrick, R., & Johnson, M. 2004. Effects of noun phrase type on sentence complexity. *Journal of memory and Language*, 51(1): 97-114.

Hamilton, H., & Deese, J. 1971. Comprehensibility and subject-verb relations in complex sentences. *Journal of Verbal Learning and Verbal Behavior*, 10(2): 163-170.

Han, C., Elouazizi, N., Galeano, C., Görgülü, E., Hedberg, N., Hinnell, J., Kim, M., Kyeong-min, & Kirby, S. 2012. Processing strategies and resumptive pronouns in English. In *Proceedings of the 30th west coast conference on formal linguistics*, 153-161. Somerville, MA: Cascadilla Proceedings Project.

Keshev, M. 2016. Active dependency formation in syntactic islands: Evidence from Hebrew sentence processing. *M.Sc. Thesis, Tel Aviv University.*

Keshev, M., & Meltzer-Asscher, A. 2017. Active dependency formation in islands: How grammatical resumption affects sentence processing. *Language*, 93(3): 549-568.

Kuznetsova, T. A., Brockhoff, P. B., & Christensen, R. H. B. 2014. lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). 2.0-6.

Lewis, R. 1996. Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of psycholinguistic research*, 25(1): 93-115.

Lewis, R. 2000. Specifying architectures for language processing: Process, control, and memory in parsing and interpretation. In M. Crocker, M. Pickering, & C. Clifton Jr. (Eds.), *Architectures and mechanisms for language processing,* 56–89. Cambridge, England: Cambridge University Press.

Lewis, R., & Vasishth, S. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive science*, 29(3): 375-419.

Mallinson, G., & Blake, B. J. 1981. *Language typology: Cross-linguistic studies in syntax.* North-Holland.

McElree, B., Foraker, S., & Dyer, L. 2003. Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, 48: 67-91.

McCloskey, J. 2006. *Resumption*, in M. Everaert and H. van Riemsdijk (Eds.), The Blackwell Companion to Syntax: 94– 117, Oxford: Blackwell.

Meltzer-Asscher, A., Fadlon, J., Goldstein, K., & Holan, A., 2015. Direct object resumption in Hebrew: How modality of presentation and relative clause position affect acceptability. *Lingua,* 166: 65-79.

Ross, J. 1967. Constraints on Variables in Syntax. *Ph.D. dissertation, MIT.*

Sells, P. 1984. Syntax and Semantics of Resumptive Pronouns. *Ph.D. dissertation, University of Massachusetts.*

Stolz, W. S. 1967. A study of the ability to decode grammatically novel sentences. *Journal of Verbal Learning and Verbal Behavior*, 6(6): 867-873.

Schlezinger, I. M. 1975. Why a sentence in which a sentence in which a sentence is embedded is embedded is difficult. *International Journal of Psycholinguistics*, 53–66.

Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. 2010. Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes*, 25(4): 533-567.

Villata, S., Tabor, W., & Franck, J. 2018. Encoding and retrieval interference in sentence comprehension: Evidence from agreement. *Frontiers in psychology*, *9*, 2.

Wagers, M., Lau, E., & Phillips, C. 2009. Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61:206-237.

Wagers, M., & Phillips, C. 2014. Going the distance: memory and control processes in active dependency construction. *The Quarterly Journal of Experimental Psychology*, 67: 1274-1304.

Warren, T., & Gibson, E. 1999. The effects of discourse status on intuitive complexity: Implications for quantifying distance in a locality-based theory of linguistic complexity. *Poster presented at the Twelfth CUNY Sentence Processing Conference, New York.*

Yngve, V. H. 1960. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society* 104: 444–466.

Appendixes

Appendix A – Experiment 1 materials

Experimental sentences

Set 1
הבלון שהליצן שהילד עצבן ניפח התפוצץ
הבלון שהליצן שהילד אותו עצבן ניפח אותו התפוצץ
הבלון שהליצנית שהילדים עצבנו ניפחה התפוצץ
הבלון שהליצנית שהילדים עצבנו אותה ניפחה אותו התפוצץ

Set 2
הפלאפון שהלקוח שהמוכר רימה קנה התקלקל
הפלאפון שהלקוח שהמוכר רימה אותו קנה אותו התקלקל
הפלאפון שהלקוחה שהמוכרים רימו קנתה התקלקל
הפלאפון שהלקוחה שהמוכרים רימו אותה קנתה אותו התקלקל

Set 3
הבניין שהאדריכל שהקונה שכר תכנן התמוטט
הבניין שהאדריכל שהקונה שכר אותו תכנן אותו התמוטט
הבניין שהאדריכלית שהקונים שכרו תכננה התמוטט
הבניין שהאדריכלית שהקונים שכרו אותה תכננה אותו התמוטט

Set 4
היהלום שהפושע שהבלש תפס גנב אותר
היהלום שהפושע שהבלש תפס אותו גנב אותו אותר
היהלום שהפושעת שהבלשים תפסו גנבה אותר
היהלום שהפושעת שהבלשים תפסו אותה גנבה אותו אותר

Set 5
הצעצוע שהתינוק שהרופא בדק קיבל התפרק
הצעצוע שהתינוק שהרופא בדק אותו קיבל אותו התפרק
הצעצוע שהתינוקת שהרופאים בדקו קיבלה התפרק
הצעצוע שהתינוקת שהרופאים בדקו אותה קיבלה אותו התפרק

Set 6
הסרט שהשחקן שהמפיק פיטר השמיץ נגנז
הסרט שהשחקן שהמפיק פיטר אותו השמיץ אותו נגנז
הסרט שהשחקנית שהמפיקים פיטרו השמיצה נגנז
הסרט שהשחקנית שהמפיקים פיטרו אותה השמיצה אותו נגנז

Set 7
הרכב שהנהג שהפקח תפס חיפש נגרר
הרכב שהנהג שהפקח תפס אותו חיפש אותו נגרר
הרכב שהנהגת שהפקחים תפסו חיפשה נגרר
הרכב שהנהגת שהפקחים תפסו אותה חיפשה אותו נגרר

Set 8
הערעור שהנבחן שהמשגיח עצבן הגיש התקבל
הערעור שהנבחן שהמשגיח עצבן אותו הגיש אותו התקבל
הערעור שהנבחנת שהמשגיחים עצבנו הגישה התקבל
הערעור שהנבחנת שהמשגיחים עצבנו אותה הגישה אותו התקבל

Filler sentences

השיר שהרשמתי את האמרגן ששמע אותו התנגן ברקע
הרמקול שהכרתי את הקריין שהחזיק אותו הועבר לזמר
הכיסא שפגשתי את הנגר שבנה אותו נמכר לבחור
הספר שהערצתי את הסופר שכתב אותו הוחזר לספרייה
המיטה שהבהלתי את המובילים שהעבירו אותה למחסן נשרטה
המעיל שהבכתי את הדוגמנית שלבשה אותו נקרע בשרוול
המאמר שפגשתי את החוקרת שכתבה אותו התפרסם בג'ורנל
הנשק שהענשתי את החיילת ששכחה אותו נמצא במלתחות
הקפה ששתיתי עם הבחור שהקופאי הרגיז נשפך
הכדור שקיבלתי מהכדורגלן שהמאמן שיבח התפוצץ
המסמך שהגשתי לפקיד שישב בדלפק נעלם
המדף שתליתי אצל הדייר שהשכנה שנאה נפל
המתנה שקיבלת ימהדוד שהההורים הזמינו הוחלפה
הפירות שבחרתי אצל הירקן שהדיאטנית אוהבת נאכלו
היצירה שהצגתי לסוחר שהגלריה שלחה נמכרה
המצגת שהכנתי עם העובד שהמנהלים חיבבו בוטלה
הצעצוע שהפעוט שאיבד אותו בכה התגלגל אל השיחים
הנעל שהכלב שנשך אותה ברח נהרסה
הטרקטור שהחקלאי שרחץ אותו הלך חנה ליד הפרדס
האוכל שהסועד שהזמין אותו התעצבן התעכב במטבח
המכתב שהמזכירה שניסחה אותו פוטרה הגיע לנמענים
המחזה שהמבקרת שאהבה אותו שיבחה הוצג בקאמרי
הצלחת שהמלצר שהחזיק אותה מעד התנפצה על הרצפה
הכרית שהחתול איתה שיחק קרע לכלבה את הסלון

Appendix B – Experiment 1 instructions

⋮   ⌄

שאלון דירוג משפטים

בשאלון זה תתבקשו לדרג משפטים על סקאלה בין 1 ל- 7 בהתבסס על עד כמה המשפט קל/קשה להבנה.

משמעות הדירוג 1 היא שהמשפט לגמרי בלתי ניתן להבנה.
משמעות הדירוג 7 היא שהמשפט ניתן להבנה בקלות.

לדוגמה, המשפט "האלבום שהלהקה שחררה לפני שנה היה מוצלח" יחסית קל להבנה. לעומת זאת, המשפט "הגלידה שהסבתא שחיכתה לנכדה החזיקה נמסה" יותר קשה להבנה.

תודה!

Appendix C – Experiment 2 materials

Experimental sentences

Set 1
הילד שהשכן שהאורח הבהיל חיבב נפל
הילד שהשכן שהאורח הבהיל אותו חיבב אותו נפל
הילד שהשכנים שהאורחת הבהילה חיבבו נפל
הילד שהשכנים שהאורחת הבהילה אותם חיבבו אותו נפל

Set 2
האיש שהחוקר שהעיתונאי העריץ בדק צחק
האיש שהחוקר שהעיתונאי העריץ אותו בדק צחק
האיש שהחוקרים שהעיתונאית העריצה בדקו צחק
האיש שהחוקרים שהעיתונאית העריצה אותם בדקו צחק

Set 3
הגנב שהמוביל שהאדריכל הכיר הפתיע נכלא
הגנב שהמוביל שהאדריכל הכיר אותו הפתיע נכלא
הגנב שהמובילים שהאדריכלית הכירה הפתיעו נכלא
הגנב שהמובילים שהאדריכלית הכירה אותם הפתיעו אותו נכלא

Set 4
התלמיד שהבמאי שהמנהל שכר העדיף פרש
התלמיד שהבמאי שהמנהל שכר אותו העדיף אותו פרש
התלמיד שהבמאים שהמנהלת שכרה העדיפו פרש
התלמיד שהבמאים שהמנהלת שכרה אותם העדיפו אותו פרש

Set 5
הטכנאי שהעוזר שהשגריר העליב מצא נשרט
הטכנאי שהעוזר שהשגריר העליב אותו מצא אותו נשרט
הטכנאי שהעוזרים שהשגרירה העליבה מצאו נשרט
הטכנאי שהעוזרים שהשגרירה העליבה אותם מצאו אותו נשרט

Set 6
הדוגמן שהשדרן שהמלצר השמיץ זיהה התפרסם
הדוגמן שהשדרן שהמלצר השמיץ אותו זיהה אותו התפרסם
הדוגמן שהשדרנים שהמלצרית השמיצה זיהו התפרסם
הדוגמן שהשדרנים שהמלצרית השמיצה אותם זיהו אותו התפרסם

Set 7
התייר שהסטודנט שהמוכר עצבן ליווה נרדם
התייר שהסטודנט שהמוכר עצבן אותו ליווה אותו נרדם
התייר שהסטודנטים שהמוכרת עצבנה ליוו נרדם
התייר שהסטודנטים שהמוכרת עצבנה אותם ליוו אותו נרדם

Set 8
הבחור שהמדריך שהדייל עיכב הצחיק נרגע
הבחור שהמדריך שהדייל עיכב אותו הצחיק אותו נרגע
הבחור שהמדריכים שהדיילת עיכבה הצחיקו נרגע
הבחור שהמדריכים שהדיילת עיכבה אותם הצחיקו אותו נרגע

Filler sentences

השיר שהרשמתי את האמרגן ששמע אותו התנגן ברקע

הנשק שהעננשתי את החייל ששבח אותו אותר במלתחות

הציור שהערכתי את האמן שצייר אותו נמכר

המיטה שהרגזתי את הסבלים שהעבירו אותה התפרקה

הסדרה שתיעבתי את התסריטאים שכתבו אותה בוטלה במפתיע

העוגה שאהבתי את המתנדבים שהביאו אותה התקלקלה אחרי יומיים

הטפסים ששנאתי את המזכירה שהדפיסה אותם אבדו

המאמר שפגשתי את המדעניות שכתבו אותו נגנז

הצעצוע שהתינוק שאיבד אותו בכה התגלגל לתוך בור

הכיסא שהקצין שרצה אותו הודח הועבר למזכירה

הבקבוק שהכלב שלעס אותו נבח נהרס

האוכל שהלקוח שהזמין אותו איחר התקרר

המכתב שהמתמחים שניסחו אותו עזבו את החברה הגיע לנמענים

ההצגה שהמבקר ששיבח אותה הוחלף הוצגה בברודויי

הכוסות שהברמן שהחזיק אותן מעד התנפצו

המפגשים שהפסיכולוגית שארגנה אותם חלתה נדחו

המתלמד שהבכיר שהכשיר אותו קודם מונה לתפקיד

האספן שהמתחזה שרימה אותו נאסר התראיין לעיתון

הפושע שהבלש שתפס אותו הודח שוחרר

הכוריאוגרף שהרקדן שהכפיש אותו נפצע שמח

הפקידים שהמפקחת שבחנה אותם הקפידה על הנהלים נכשלו בביקורת

החשוד שהפרקליטה שחקרה אותו ניצחה במשפט נכנס לכלא

הפקח שהטייסים שהדאיגו אותו יצרו קשר התעודד

הפועלים שהקבלן שקָנַס אותם זכה במכרז מחו

ניסוי הבנת משפטים

שלום!

בניסוי זה תקרא.י משפטים על מסך המחשב. כל משפט יופיע מילה-אחר-מילה במרכז המסך. המילים יופיעו בקצב קבוע ולא תוכל.י לחזור לקרוא אותן לאחר שיעלמו.

מיד לאחר שתסיימ.י לקרוא את המשפט, תתבקש.י לענות על שאלת הבנה הנוגעת למשפט שאותו קראת. בפניך יוצגו שתי אופציות לתשובה, השתמש.י בעכבר על מנת לבחור את זו המתאימה לדעתך.

על מנת לעבור ממשפט למשפט הבא יש ללחוץ על מקש כלשהו במקלדת.

השתדל.י לשמור על ריכוז מירבי על מנת שתצליח.י לקרוא את המשפטים ולבצע את המטלה. כמו כן, נבקש שתבצע.י את הניסוי במקום שקט, ברצף וללא הפרעות.

כשתסיימ.י לעבור על ההוראות, תועבר.י למספר **משפטי אימון** ואחריהם יתחיל הניסוי עצמו.

תקציר

משפטי סנטר אמבדינג (center embedding, שעבוד מרכזי), כדוגמת 'הסלמון שהאיש שהכלב נשך עישן היה טעים', שמכילים שתי פסוקיות זיקת מושא המשועבדות אחת תחת השנייה, הינם ידועים לשמצה מבחינת קושי העיבוד שעולה בהם (חומסקי ומילר, 1963; בלטין וקולינס, 2008). שני הסברים עיקריים הוצעו באשר לקושי העיבוד במשפטים אלה. גיבסון (1998) טוען שהוא נובע מעלויות אחזקה ואינטגרציה: עלויות האינטגרציה על הפועל השני גבוהות ממה שקיבולת זכרון העבודה של רוב האנשים יכולה להכיל, ולכן הם נכשלים בשיוך הפילרים (fillers) לפעלים המתאימים להם. לעומת זאת, לואיס ווסישת' (2005) טוענים שהקושי במבנים אלה עולה בזמן השליפה: בהיעדר רמזי שליפה מספיקים, השליפה של הפילר בזמן ההגעה לפועל נכשלת בשל הדמיון בין שלושת הצירופים השמניים, מה שמוביל לאינטרפירנס (interference, הפרעה).

המחקר הנוכחי מתמקד במשפטי סנטר אמבדינג בעברית ובוחן האם רמת המובנות וההבנה בפועל שלהם יכולות להיתרם מנוכחות של: (i) תכוניות התאם שמבחינות בין שלושת הצירופים השמניים ומזהות את הנושא של כל פועל, ו- (ii) כינויי גוף חוזרים (שהינם דקדוקיים ומופיעים בתפוצה חופשית באופן יחסי בעברית), שיכולים לתרום לשליפה בכך שהם מספקים זמן עיבוד נוסף ו/או בכך שהם מסומנים בתכוניות ההתאם של הפילרים ולכן מאפשרים זיהוי חד משמעי של המושא של הפועל.

ניסוי 1 (160 משתתפים) דן בשאלה זו בעזרת מטלת רמת הבנתיות. הוא כלל ארבעה תנאים שהורכבו מהצלבה בין הגורמים *התאם מובחן* (תכוניות ההתאם על שלושת הצירופים השמניים זהות או שונות) ו*שימוש בכינוי גוף חוזר* (המושאים של הפועל מומשו ע"י גאפ (gap) או כינוי גוף חוזר). המשתתפים בניסוי קראו את המשפטים בקצב חופשי וללא הגבלת זמן ודירגו את רמת ההבנתיות שלהם על סולם של 1-7. התוצאות הראו שהן *התאם מובחן* והן *שימוש בכינוי גוף חוזר* לא השפיעו באופן מובהק על רמת ההבנתיות של המשפטים. נצפתה אינטרקציה (interaction) מובהקת בין שני הגורמים (p=.03), שהעידה על תרומה של התאם מובחן רק בהיעדר כינוי גוף חוזר.

ניסוי 2 (192 משתתפים) עשה שימוש במטלת שאלות הבנה שהופיעו לאחר המשפטים. המשפטים הניסויים הורכבו מאותם ארבעה תנאים של ניסוי 1. בשאלות ההבנה נבדק גם הגורם *שאלת פועל*, שנשאלה או על המושא של הפועל הראשון או על המושא של הפועל השני. המשפטים הוצגו מילה-אחר-מילה בקצב של 400 מילישניות לכל מילה + 200 מילישניות הפסקה בין מילה למילה. התוצאות הראו שה*התאם מובחן* שיפר את ההבנה של המשפטים באופן מובהק (p=.004), בעוד ש*שימוש בכינוי גוף חוזר* לא. האינטרקציה בין הגורמים הללו לא הייתה מובהקת, מה שמעיד על כך שלא נצפה הממצא של ביטול התרומה של התאם מובחן במקרים של שימוש בכינוי גוף חוזר, בניגוד לניסוי 1. התוצאות הראו גם אפקט (effect) *של שאלת פועל* (p=.001), כך שהפועל המשועבד ביותר (וסגירת התלות שלו) היה הקשה ביותר. האינטרקציה בין שאלת פועל להתאם

מובחן הייתה מובהקת (p=.001) והראתה שבעוד *שלהתאם מובחן* לא הייתה תרומה לסגירת התלות של הפועל המשועבד ביותר, וכפועל יוצא מכך להבנה שלו, הייתה לו תרומה להבנה של הפועל השני.

התוצאות של ניסוי 2 מעידות שמשפטי סנטר אמבדינג ניתנים להבנה ברמה זו או אחרת, בייחוד בהינתן התאם מובחן על הצירופים השמניים. לעומת זאת, כינויי גוף חוזרים, למרות הפוטנציאל שלהם לאפשר זיהוי חד משמעי של המושא של כל פועל, לא תרמו להבנה. התוצאות האלה עשויות להעיד על כך שבכינויי גוף חוזרים לא משמשים לצרכי שליפה, או שאינטרפירנס התרחש כבר בשלב הקידוד של שלושת הצירופים השמניים הדומים (גורדון, הנדריק וג'ונסון, 2004; וילאטה, טאבור ופראנק, 2018), וגרם לכך שהפילרים לא היו מובחנים מספיק כדי לאפשר שליפה מוצלחת בעת ההגעה לפועל. שימוש בכינוי גוף חוזר לא תרם להבנה, ויותר מכך הוא ביטל את התרומה של התאם מובחן בניסוי 1. ניתן להסביר את הממצא הזה בדומה להסבר של "the missing V2 effect" (אפקט הפועל השני החסר), ההבחנה שמשפטי סנטר אמבדינג נשפטים כקבילים יותר כאשר רק שניים מהפעלים מופיעים במשפט (פרייז'ר, 1985; גיבסון ותומאס, 1999). גיבסון ותומאס מציעים שבמקרים אלו אחת התלויות נשכחת והקושי העיבודי של המשפט מטשטש. בהישען על ההסבר הזה, ניתן להניח שהשימוש בכינוי גוף חוזר לא מאפשר להזניח את אחת התלויות ומוביל לדירוגי הבנתיות נמוכים.

אוניברסיטת תל-אביב

הפקולטה למדעי הרוח ע"ש לסטר וסאלי אנטין

החוג לבלשנות

הנושא:

שלושה צירופים שמניים נכנסים למשפט: האם התאם ושימוש בבינויי גוף חוזרים יכולים לעזור לשפר משפטי סנטר אמבדינג?

חיבור זה הוגש כעבודת גמר לקראת התואר

"מוסמך אוניברסיטה" - M.A. באוניברסיטת ת"א

על ידי

הילה דוידוביץ

העבודה הוכנה בהדרכת:

פרופ' איה מלצר-אשר

תאריך:

אוקטובר 2021