Now you see *her*; now you don't: A visual world eye-tracking study of pronoun resolution in real time

Rachel Gargiulo

1st of August, 2015

Research Master's in Linguistics

Utrecht Institute of Linguistics (UiL OTS)

Utrecht University, Utrecht, The Netherlands

Supervisors: Dr. Frank Wijnen, Dr. Iris Mulders

Abstract

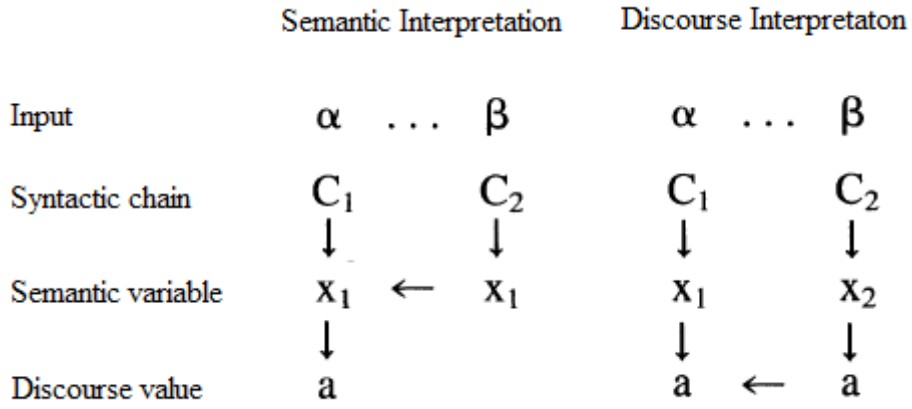
Reuland's (2001) Principles of Binding (PoB) framework proposes that there are multiple routes to comprehending anaphora. It also incorporates an *economy hierarchy* which predicts relative 'costliness' of each route according to the amount of cognitive resources required for successful anaphor resolution. Koornneef (2008) conducted a series of experiments in order to determine whether the predictions made by the PoB are instantiated in psychological reality. The results of an eye-tracking reading experiment by Koornneef (2008) offered support for the PoB framework and economy hierarchy. However, in a post-hoc analysis, Koornneef found an effect of individual differences that seem antithetical to the principle of economy. Some participants appeared to voluntarily calculate the more costly route. One alternative explanation may lie in working memory capacity. Pronouns are inherently ambiguous, and some studies suggest that availability of working memory resources is a factor in resolving ambiguity. I replicated the experiment by Koornneef (2008) to investigate whether this could explain his findings, but used the visual world paradigm in order to determine whether looking behavior can track pronoun resolution processes in real time, potentially leading to the ability to detect whether working memory capacity plays a significant role in pronoun resolution. Participants listened to short stories about two characters while looking at a visual display and their eye-movements were recorded. These stories contained two regions of interest: an overt pronoun and an elided verb phrase containing a pronoun. At the overt pronoun, participants attended more strongly to the character representing the non-antecedent when the condition contained a higher level of ambiguity compared to the low-ambiguity condition. In addition, there were differences between participants with high and low working memory capacity, indicating that this may be a factor in ambiguity resolution. However, at the ellipsis site, a strong and unexpected pattern of attention to the same character across participants and conditions suggests that the VWP is sensitive to other discourse processing factors which can create confounds and influence looking behavior, calling into question the interpretation of the findings at the overt pronoun.

1. Introduction

According to the Principles of Binding (PoB) framework (Reuland 2001), there are two linguistic levels, or modules, in which pronouns can be linked to their antecedents: semantics and discourse (see also Heim 1998). An important feature incorporated into the PoB framework is its *economy hierarchy*. The principle of economy stems from the limited nature of cognitive processing resources. Since we have finite resources, the cognitive system in general, and the language system by

extension, tends to minimize working memory costs (Reinhart 1983). In PoB, the principle of economy is applied to pronoun resolution by "counting interpretive steps" (Reuland 2001:485). Figure 1 below shows how the steps proceed. In this illustration, *α* and *β* represent two input expressions; for example, a pronoun and its antecedent. *C* represents a syntactic representation, or chain[1]. *x* represents the variables that are stored in the semantic representation. *a* represents the value as stored in the discourse representation.

Fig. 1. Based on Reuland's (2001) illustration of the cross-modular interpretive steps required for pronoun resolution at the semantic and discourse levels.



Moving information from one level to another (e.g., from the semantic module to the discourse module) counts as a cross-modular step. Processing costs are incurred with each cross-modular step, meaning that the extra step required in carrying over the values of α and β separately to the discourse representation makes pronoun resolution in the discourse module costlier than in the semantic module.

The method by which the semantic module encodes this referential dependency is known as *variable binding* (1a), while the discourse operation is known as *coreference* (1b).

(1) *The policeman*$_a$ knows that *he* is guilty.

a. *x* knows *x* is guilty

b. *x* knows *a* is guilty; *a* = *x*

In (1), both resolution operations end in the same interpretation: the policeman knows himself to be guilty. In a discourse environment, there are often multiple possible antecedents. Since a c-command relationship is required for variable binding to occur, and only one of the possible antecedents in (2)

[1] Syntactic chains are not relevant for our purposes here because they are formed in the syntactic representation. Here we will focus only on the semantic and discourse modules. For a full discussion see Reinhart & Reuland (1993) and Reuland (2001).

c-commands the pronoun (specifically, *the policeman*), variable binding can still only result in one interpretation (2a); coreference, however, can now lead to multiple interpretations (2b,c).

(2) The trial of *a bank robber*$_a$ begins today. *The policeman*$_b$ knows that *he* is guilty.

a. *x* knows *x* is guilty

b. *x* knows *a* is guilty

c. *x* knows *b* is guilty; *b* = *x*

In this example, world knowledge and the given discourse context biases toward the interpretation in (2b), but it is possible that a policeman robbed a bank and framed a known bank robber, in which case (2a) would represent the semantic (bound-variable) interpretation of the pronoun, and (2c) would represent the contextually appropriate discourse (coreference) interpretation. Put differently: there are two routes to assigning pronominal reference, but variable binding operates in a limited syntactic context. When variable binding fails, coreference is established within the discourse module. The economy hierarchy predicts that coreference, though available, should not be preferred to variable binding, so that (2a) will always be preferred over (2c). There are some limitations to coreference as well, but these mainly arise from cognitive and linguistic factors that either constrain or bias interpretive processes, including but not limited to recency (Clark & Sengul 1979), semantic gender information (e.g., Arnold, Eisband, Brown-Schmidt & Trueswell 2000), implicit causality information of interpersonal verbs (e.g., Koornneef & Van Berkum 2006), and first mention (e.g., McDonald & MacWhinney 1995).

In order to investigate the psychological reality of the PoB framework, Koornneef (2008) conducted an eye-tracking reading experiment. As part of a larger story, participants read sentences such as in (3). The regions of interest are in italics.

(3) Roos ziet toekomst in *haar* relatie, maar *Anouk helaas niet.*
Rose sees a future in *her* relationship, but *Anouk unfortunately doesn't*.

The second region of interest, *Anouk unfortunately doesn't* carries the underlying meaning in (4)

(4) Anouk unfortunately doesn't see a future in *her* relationship.

There is a pronoun, *her*, that is unpronounced, but understood. Half of the items in Koornneef's experimental materials were biased toward a bound-variable interpretation of the pronoun, and the other half were biased toward a coreferential interpretation.

He predicted that participants would read more quickly in stories that biased the reader toward a bound-variable interpretation of a pronoun as compared to stories that biased toward a coreferential interpretation because the bound-variable interpretation incurs fewer processing costs, according to the PoB. Overall, Koornneef's (2008) results supported the PoB-based predictions, with one caveat. In a post-hoc analysis, he found that his participants could be divided into two groups, each displaying a different behavioral pattern. While one group of participants read through the pronoun *her* slowly, relative to the other group, they were faster while reading the ellipsis *Anouk unfortunately doesn't*. The other group showed the opposite pattern. They read relatively quickly during the pronoun region, and more slowly than the other group during the ellipsis region. Koornneef (2008) hypothesized that the reason for the different reading times was that some people made the effort to compute both the bound-variable and coreferential interpretations of the overt pronoun, which then led to an advantage in processing the elided pronoun. As Koornneef notes, it is unclear why this should be the case under the PoB. The PoB model with its economy hierarchy does not predict that some people would choose to calculate the costlier discourse dependency in addition to the cheaper semantic dependency at the overt pronoun, especially since both calculations lead back to the same antecedent (*Rose sees a future in her relationship*).

Studies on processing ambiguity and working memory suggest an alternative explanation for Koornneef's (2008) pattern of results. There are two layers of ambiguity present in the regions of interest. The first is the latent ambiguity in the overt pronoun (*her* in the example above) as potentially referring to anyone in the discourse context, but this pronoun is *combinatorially unambiguous*; in other words, resolving this pronoun is straightforward as both interpretations lead to the same antecedent. The second layer is that the elided pronoun (in *...Anouk unfortunately doesn't (see a future in **her** relationship...*) is *combinatorially ambiguous* – each interpretation leads to a different antecedent – and further information from the discourse biases interpretation. Participants with lower working memory capacity either may not 'notice' or fully process the ambiguity (Nieuwland & Van Berkum 2006[2]) at the overt pronoun, which would have shown up as faster reading times in that region of interest in Koornneef's (2008) study. However, when the ellipsis is encountered, it becomes necessary to process the ambiguity, and this would result in slower reading times during that region of interest. Participants with higher working memory capacity may initially devote more resources to processing the ambiguity at the overt pronoun, which would appear to be a disadvantage in terms of reading times, but would later become an advantage in terms of reading time at the ellipsis.

I aimed to replicate Koornneef's (2008) experiment using the visual world eye-tracking paradigm (VWP) in order to further investigate his hypothesis and the alternative by observing looking behavior. My aim was to test whether working memory, as a cognitive constraint on

[2] The evidence for this is discussed in Section 2.3.

ambiguity resolution, predicts looking behavior during pronoun resolution. The ability of the VWP to record looking behavior with high temporal resolution gives it the potential to explore some of the more fleeting cognitive processes.

2. Language comprehension online

Koornneef (2008, 2010:20) argues that "the most straightforward implementation [of the PoB framework] involves a serial architecture, consisting of three distinct processing phases" closely resembling the language processing model proposed by Friederici (2002).

2.1 The PoB online

Friederici's (2002) model of the language processing architecture can be seen as a neurocognitive instantiation of the relevant assumptions of the PoB model (Koornneef 2008, 2010). Based on electrophysiological and brain-mapping data, Friederici (2002) identified three 'phases' during the comprehension process. During Phase 1 (100 – 300 ms), syntactic structure is built on the bases of lexical category information. During Phase 2 (300 – 500 ms), lexical-semantic and morphological processes are engaged to assign thematic roles. Phase 3 (500 – 1000 ms) is the integration phase, during which thematic role assignment is mapped onto syntactic structure. So, different kinds of information are processed by different modules, in different time windows, and then are afterwards integrated. Here, we will focus on the stages of comprehension that occur after perception and phonological processes.

The phases were identified by means of event-related potential (ERP) components. An early left anterior negativity (ELAN) is often seen in response to word-category violations, and occurs in the same time-window as Friederici's Phase 1. A negativity at approximately 400 ms (N400) tends to be a response to semantic anomaly, and occurs during the same time window as Phase 2. There are two ERP components that are correlated with syntactic anomaly. One is the left anterior negativity (LAN), occurring fairly early on, between 100 and 500 ms after the stimulus and in response to morphosyntactic errors. The other is a centro-parietal positivity (P600) occurring at approximately 600 ms, during Phase 3, in response to syntactic violations and while processing syntactically complex material. The serial nature of the model was motivated by the *semantic blocking* effect observed in Friederici, Steinhauer, & Frisch (1999), Hahne & Friederici (2002), and others. Hahne & Friederici (2002) conducted an electrophysiological study investigating sentences with only a syntactic violation compared to sentences with both a syntactic and a semantic violation occurring in

the same word. The latter should lead to processing difficulties in both domains. The examples in (5) are the translations given of the original German.

(5) a. The shirt was ironed. (correct)
b. The shirt was on ironed. (syntactic violation)
c. The thunderstorm was in ironed. (combined violation) [Friederici 2002:82]

They hypothesized that, in the combined violation condition, the ELAN, N400, and P600 components should all be present. Rather, they found an ELAN and a P600 but no N400. They interpreted this to mean that syntactic structure building precedes semantic processing, so, in the case of (5c), the semantic meaning of *ironed* is never processed because the lexical item was rejected at phase 1 – it was not syntactically licensed, so it was never integrated, hence the term *semantic blocking*.

The serial architecture combined with the economy hierarchy translates into pronoun interpretation is as follows. The first opportunity to establish a link between a pronoun and its antecedent is via semantics, which is also the less costly route. If this fails because there is no c-command relationship, then the variables are transferred separately to the discourse module to be valued. Alternatively, if a link is successfully established via semantics, then information from the discourse comes into play and there may or may not be a conflict between the semantic interpretation and the discourse context. If there is a conflict (i.e., the semantic interpretation has led to a contextually inappropriate antecedent), then the appropriate antecedent is coreferenced via discourse, the more costly interpretation according to the economy heirarchy because it requires the extra step of separately transferring the relevant variables into the discourse module.

2.2 Individual differences in language processing

The studies by Friederici and colleagues described in the previous section rely on grand-mean analyses of their data for their conclusions. This is, of course, the rule rather than the exception, but there is good reason to expand how we approach data analysis to include looking at individual differences. The story of how language comprehension proceeds online can appear somewhat different from this perspective. Recent studies using electroencephalography (EEG) to measure neural mechanisms related to semantic and syntactic anomalies are all but forcing a reevaluation of how we understand the various ERP components that we see in response to linguistic stimuli, and are problematic for the story as told above.

Very few (published) EEG studies have looked at individual differences in ERP response profiles. In fact, as far as I'm aware, they can be counted on one hand. Typically, only grand means are analyzed. The first study that investigated individual differences in ERP effects to linguistic

stimuli was Osterhout (1997). He found that, while a P600 effect was robust for syntactically anomalous closed-class items, individual differences can be observed for syntactically anomalous open-class items. These elicited a P600 response in most subjects, but an N400 response in others. More recently, work by Tanner (2013, 2014), Tanner & Van Hell (2014), and Nickels, Bokhari, & Steinhauer (2014) have revealed the necessity of questioning how the data from electrophysiological measurements are analyzed, and that we should be careful interpreting grand-averaged data. Tanner & Van Hell (2014) used ERPs to investigate processing of morphosyntax in two conditions: subject-verb agreement (6a) and verb tense (6b).

(6) a. The clerk at the clothing boutique was/*were severely underpaid and unhappy.
b. The crime rate was increasing/*increase despite the growing police force.

In their grand-mean analysis, they found a biphasic LAN/P600 response pattern to grammatical violations. However, when they analyzed individuals' response patterns, they found that their participants varied across a continuum, ranging from negativity-dominant to positivity-dominant responses. This individual analysis also revealed that the LAN can be a variant of the N400. Overlap between right-hemisphere topography of the P600 component and the centro-parietal N400 component was responsible for left-hemisphere topography of the negativity. These findings led the authors to conclude 1) a biphasic ERP in the grand mean does not straightforwardly indicate distinct stages of processing in individual participants, 2) "there are multiple neurocognitive routes to successful grammatical comprehension" (Tanner & Van Hell 2014:289).

Nickels et al. (2014) attempted to replicate the study by Friederici, Steinhauer & Frisch (1999), mentioned above, using identical materials[3] and adding one condition[4], in order to investigate an alternative explanation for the observed semantic blocking effect. Not only did they fail to replicate the semantic blocking effect, they had findings similar to Tanner & Van Hell (2014) in that their participants showed a range of ERP response profiles when inspected individually.

These examples point toward the need to expand our investigations of the psychological reality of language processing to include potential sources of individual differences. In particular, as Koornneef's (2008) post-hoc analysis showed, the psychological reality of the PoB framework may be accurate for many, but cognitive factors, such as differences in working memory capacity, may

[3] Much like those in example (5).
[4] The added condition was meant to test whether the absence of an N400 in the original study might be explained by participants processing the final adjective as part of a new prepositional phrase (Steinhauer and Drury 2012). (i) is the English translation of an example of a fragment that appeared in Friederici (1999), and (ii) is an example of the condition added by Nickels et al. (2014).
(i) The priest was to-the-asphalted
(ii) The priest was to-the-asphalted crossing called

influence language processing in such a way that it is not accurate or simply not relevant for everyone.

2.3 Working memory & ambiguity

Swinney (1979) showed that lexically ambiguous items initially activate multiple meanings. For example, *bug* will initially activate both insects and spying devices, even within a strongly biasing prior context. The meanings that are not contextually appropriate become less available rapidly. In the original study, this occurred at approximately 750 to 1000 ms. However, Swinney points out that, in normal processing, it likely occurs much more rapidly – a matter of perhaps 300 milliseconds or so (Swinney 1979; Swinney, Prather & Love 2000). A study by Gernsbacher (1989) indicates that a similar process may occur with pronouns, but more locally-contextually driven rather than lexically driven, so that only the relevant referents of the current discourse context are activated. Gernsbacher (1989) looked at what she termed 'suppression' versus 'enhancement' of antecedents. Enhancement increases activation of mental representations while suppression decreases activation. Gernsbacher demonstrated, using a probe-verification task, that while a repeated name activated only one out of two possible antecedents, a (singular 3rd person) pronoun activated both antecedents (7a), and this occurred even when the pronoun followed a disambiguating phrase, as in (7b).

(7) a. Ann predicted that Pam would lose the track race, but *Pam/she* came in first very easily.

b. Ann predicted that Pam would lose the track race. But after winning that race *she*…

One way of operationalizing working memory capacity is by testing *reading span*[5] (Daneman & Carpenter 1980). Nieuwland & Van Berkum (2006) looked into correlations between reading span and resolving ambiguous pronouns, and found individual differences in ERP responses. The ambiguous pronouns elicited a sustained frontal negativity (Nref), but the size of the effect varied across participants, being positively correlated with reading span scores. Those who scored higher on the reading span task showed an Nref that was larger in amplitude. The authors suggest that high-span individuals are either more sensitive to the multiple possible interpretations of ambiguous pronouns, or are more able to pursue both interpretations due to having more processing resources. The subtle difference between these two possibilities lies in whether the low-span individuals 'notice' the ambiguity in the pronoun but don't have the resources to pursue both interpretations, or whether they settle on one interpretation without noticing the ambiguity. Nieuwland (2014) also used EEG to look at effects of gender mismatch between pronouns and their antecedents. Mismatching pronouns elicited

[5] The reading span task is described fully in *Methods*, Section 4.2.1.

an Nref compared to matching pronouns. However, this effect was robust only in participants with high reading span scores. Those with low reading span scores instead showed a P600 effect. Nieuwland suggested that this is consistent with an attempt at coreferential interpretation in spite of the mismatching gender. This means that low-span readers are less likely to read *she* in example (8) as referring to a person unmentioned in the discourse.

(8) John shouted that she was very angry today.

The results from both of these studies accord with data from St. George, Mannes & Hoffman (1997), which indicates that people with low reading span scores are also less likely to elaborate their situation model or discourse representation with optional, knowledge-based inferences. In Nieuwland's study, this is exactly what would be required to posit a discourse-external 2nd character for a sentence like (8) if encountered without any other context.

A number of studies provide further evidence that working memory resources can cause differences in how people process pronouns (e.g., Vasić 2006; Vasić, Avrutin & Ruigendijk 2006; Shapiro, Hestvik, Lesan & Garcia 2003). Stronger preferences for bound-variable readings of pronouns are associated with lower working memory resources. Children and aphasics are among those with strong bound-variable preferences. While a preference does exist in unimpaired adults, it has been found to be weaker. Working memory capacity appears to modulate the relevance of the PoB's economy hierarchy. The lower a person's working memory capacity is, the more psychologically 'real' the economy hierarchy becomes.

How this fits into the context of Koornneef's (2008) findings may seem counterintuitive at first. I propose that the faster readers during the overt pronoun were those with lower working memory capacity. This is based on two factors: (i) recall the suggestion made by Nieuwland & Van Berkum (2006) that higher working memory capacity means an increased ability to 'notice' and/or process multiple possible interpretations of a pronoun. This may result in slightly longer reading times for those who notice/process the existence of multiple possibilities than for those who do not; (ii) the overt pronoun is, in the case of Koornneef's (2008) study as discussed above, *combinatorially unambiguous*, so that any underlying natural ambiguity is unimportant and the discourse context requires no further evaluation, making the underlying ambiguity easily dismissed or ignored if resources are limited.

The discussion above describes one explanation, based on working memory capacity, for Koornneef's (2008) observation that some of his participants read more slowly through the overt pronoun than others, but I have yet to address how working memory makes the opposite prediction at the ellipsis site for the same set of participants; that is, participants with higher working memory capacity should read more quickly through the ellipsis site even though they read more slowly at the

overt pronoun. This proposal is based on (i) the evidence discussed above that higher working memory capacity can increase the ability to cope with ambiguity; (ii) a pronoun's ambiguity is increased within an ellipsis site, and cannot be dismissed, slowing down readers with lower working memory. This requires some discussion of verb-phrase ellipsis.

3. VP-ellipsis

The term *ellipsis* is used in linguistics to describe the phenomenon of material being unspoken, but comprehension of this unspoken material is nevertheless necessary for full interpretation. Verb-phrase (VP) ellipsis is the name for the phenomenon exemplified in (9), whereby ellipsis is applied to a verb phrase[6].

(9) Sam Harris thinks he is brilliant but Chomsky doesn't. (…think he is brilliant.)
   a. Chomsky doesn't think of himself as brilliant.
   b. Chomsky doesn't think Sam Harris is brilliant.

This example also shows that the elided pronoun is ambiguous. (9a) represents the bound-variable interpretation and (9b) the coreferential interpretation.[7]

A large body of experimental studies has led to alternative views on the exact nature of both the antecedent of the elided material and of the content of the ellipsis site. The nature of the antecedent (a VP in the case of VP-ellipsis, such as *thinks he is brilliant* in example (9)) may be either a semantic/discourse entity (e.g., Hardt 1993) or a syntactic entity (e.g., Fiengo & May 1994), while the content of the ellipsis site is said by some to contain a detailed structure (e.g., Fiengo & May 1994), and by others to be a cue, such as a pointer or anaphor (e.g., Martin & McElree 2008). Since the purpose of the current study is to replicate an experiment in Koornneef (2008), it seems appropriate to maintain his assumptions in order to enable more direct comparison of results. This guides the following description of how ellipses are interpreted (for a full review and discussion, see Phillips & Parker 2013).

[6] What is VP-ellipsis in English is argued to be a slightly different construction in Dutch – bare argument ellipsis (Reinhart 1991). However, the important factor for current purposes is that the ellipsis creates an ambiguity between bound-variable and coreferential interpretations of a pronoun. This is true for both languages.

[7] Though it was pointed out in Section 1 above that coreference can, in fact, lead to both antecedents, the economy principle holds that the less costly option should be preferred. Thus, coreference should only be used if the antecedent accessed by the bound-variable interpretation is contextually inappropriate. For expository ease, I henceforth refer to the antecedent that is unavailable to the variable-binding operation as *the* coreferential interpretation, even though it's one of two possibilities.

Simply put, the explicit assumption in Koornneef's (2008, see pp. 129-30) study was that interpretation of an ellipsis is guided by a copy-and-paste procedure, whereby a phonologically null copy of the antecedent VP is inserted into the ellipsis site. Following Vasić (2006) he proposed that two copies of the antecedent are made: one containing the bound-variable dependency, and one containing the discourse dependency. If some participants calculated only the bound-variable dependency at the overt pronoun, then this would be copied into the ellipsis site, so that a bound-variable interpretation at the overt pronoun leads to a bound-variable interpretation at the ellipsis. If the context biases toward a coreferential reading at the ellipsis, then there is a small delay as the discourse dependency is calculated. Those participants who calculated both the bound and discourse dependencies at the overt pronoun are at an advantage because they can simply choose which dependency to copy-and-paste for the contextually appropriate interpretation.

The problem this hypothesis faces is that it is antithetical to the economy hierarchy, which Koornneef's results otherwise support. If the answer lies in the elevated levels of ambiguity present in pronouns within VP-ellipsis sites, then the finding is in line with the economy hierarchy as modulated by working memory capacity.

4. Experiment: Visual world paradigm

4.1 The original experiment

According to the PoB's economy hierarchy (Reuland 2001), there are fewer processing costs incurred by variable binding compared to coreference. Based on this, Koornneef (2008) predicted that participants should prefer a bound-variable reading over a coreference reading of a pronoun, regardless of the preceding discourse, leading them to initially construct the 'wrong' (bound-variable) representation. This should lead to longer reading times and more backtracking to previously-read material (Koornneef 2008:115) while calculating the coreferential interpretation compared to when the story is biased toward a bound-variable interpretation. Koornneef (2008) conducted an eye-tracking reading experiment to investigate these predictions.

The examples below show the two conditions used in the study. (10) represents the bound variable condition. The critical manipulation is underlined, and the two regions of interest – the possessive pronoun and the ellipsis – are italicized.

(10) *Bound-variable condition:*

Luuk and Stijn volgen de cursus 'Koken voor dommerdjes'. Vandaag moesten zij een appeltaart bakken en hun projectjes kan men het best omschrijven als 'interessant'. Luuk durft *zijn* creatie zelfs niet te proeven, maar *Stij wel*__. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus, maar je weet maar nooit.

English translation:
Luuk and Stijn follow the course 'Cooking for dummies'. Today they had to bake an apple pie and their little projects can best be described as 'interesting'. Luuk dare not even taste *his* creation, but *Stijn does*__. So far there are no deaths during the course, but you never know.

In this example, the unpronounced pronoun within the ellipsis site is ambiguous, but should be biased toward taking *Stijn* as an antecedent, because *Stijn* represents the less costly interpretation. In other words, the full interpretation could be either *Stijn dares to taste Luuk's creation* or *Stijn dares to taste his own creation.* In the coreference condition (11), the elided pronoun is biased toward taking *Luuk* as the antecedent, since he is the only character who has baked an apple pie. Readers were predicted to slow down in the ellipsis site in the coreference condition compared to the bound-variable condition as a result of the extra processing costs incurred by the need to calculate coreference.

(11) *Coreference condition:*
Luuk and Stijn volgen de cursus 'Koken voor dommerdjes'. Vandaag moest Luuk een appeltaart bakken en zijn projectje kan men het best omschrijven als 'interessant'. Luuk durft *zijn* creatie zelfs niet te proeven, maar *Stijn wel*__. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus, maar je weet maar nooit.

English translation:
Luuk and Stijn follow the course 'Cooking for dummies'. Today Luuk had to bake an apple pie and his little project can best be described as 'interesting'. Luuk dare not even taste *his* creation, but *Stijn does*__. So far there are no deaths during the course, but you never know.

Indeed, the total time spent reading and the likelihood of backtracking were longer in the coreference condition. However, when Koornneef looked at individual differences, he found that he was looking at two groups of participants. In the coreference condition, one group read more quickly (in comparison to the other group) through the overt pronoun in *Luuk durft* ***zijn*** *creatie…* and more slowly during the ellipsis region *Stijn wel*__. The other group read relatively slowly through the pronoun, but more quickly at the ellipsis site. He suggested that those who read more quickly at the overt pronoun did so because they only calculate one interpretation of the pronoun, and this leads to a

disadvantage during the ellipsis. On the other hand, the group who reads more slowly does so because they calculate both the coreferential and the bound-variable interpretations of the possessive pronoun *zijn*, though they yield the same result. This leads to an advantage at the ellipsis site. The explanation for the advantage is that the calculations have already been made and can simply be copied and pasted in order to resolve the ellipsis. The readers who did not take the trouble to calculate both have to do the work when they reach the ellipsis. He further suggests that this is subject to strategizing or perhaps even conscious control. The trouble with this explanation, as Koornneef also acknowledges, is that it's not at all obvious why, if there is a clear preference for bound-variable interpretations in general, and a bias towards variable binding in example (10) above, anyone would bother to compute the more costly coreferential interpretation.

4.2 The current experiment

The findings of Gernsbacher (1989), Nieuwland & Van Berkum (2006), Nieuwland (2014), and Swinney (1979) converge to suggest an alternative explanation. Swinney (1979) found that multiple meanings of lexically ambiguous words are activated immediately and automatically during lexical access. The contextually irrelevant meanings are then rapidly deactivated. Gernsbacher (1989) showed that pronouns initially activate multiple antecedents as well, regardless of disambiguating context, suggesting that this may also be automatic. In the case of the overt pronoun in the experimental items used by Koornneef (2008) and in the current study, if it has two possible antecedents then we can expect both to be activated simultaneously, and one rapidly deactivated, or suppressed in Gernsbacher's terminology. Nieuwland & Van Berkum (2006) and Nieuwland (2014) found that individuals with higher reading span scores appear to have an increased ability to either 'notice' the ambiguity in pronouns or pursue multiple possible interpretations compared to individuals with low reading span scores. This fits Koornneef's (2008) pattern of results. If this *pattern* of findings can be replicated, and correlates with reading span scores, then we have a possible answer as to why Koornneef's results appear antithetical to the economy hierarchy. Individuals with higher working memory capacity may not need to conserve resources as much as those with lower working memory capacity. In other words, the economy hierarchy may be more or less relevant depending on working memory capacity. The current experiment aimed to test this hypothesis.

The current study made use of the same experimental materials as in the original, but exploited the visual world paradigm (VWP) with a 'look and listen' task. One advantage of this paradigm is that it relies on the well-attested tendency of people to look at relevant images as they are referred to without involving the potential confounds or behavioral changes that might result from task demands in, for example, probe-recognition or lexical decision tasks (Huettig, Rommers & Meyer 2011). It also allows looking behavior, reflective of attention, to be tracked with high temporal

resolution. Another advantage is that the VWP has been used several times in past studies to investigate pronoun resolution (e.g., Arnold et al. 2000; Arnold 2001; Cozijn et al. 2011; Kaiser et al. 2009) and has shown itself to be sensitive to a number of relevant constraints, such as word order, information structure, and syntactic role (Huettig et al. 2011). My aim was to test whether the VWP is also sensitive to the time-course of pronoun resolution, and whether this might also be extended to pronoun resolution in VP-ellipsis.

4.2.1 Method

Participants. Thirty-two students were recruited from the Utrecht University community, and were paid for their participation. 26 female, 6 male, between 18 and 33 years of age, mean age 22. All were native speakers of Dutch, and had normal hearing and normal or corrected-to-normal vision. The data from 2 participants was excluded from analysis due to excessive blinking or staring. Thus, 30 participants were included in the analysis.

Materials. The auditory stimuli used were originally constructed and pre-tested by Arnout Koornneef (see Koornneef 2008, p. 107), and were used with his permission. Experimental items were short stories, in Dutch, consisting of four sentences each, and occurring in two conditions: bound-variable biased (12) and coreference biased (13). The first sentence introduced two characters and a situation. The second sentence contains either a plural pronoun referring to both characters (bound-variable biased), or the name of one character (coreference biased). The third sentence contains a singular pronoun, and ends in an ellipsis. The fourth sentence provides a wrap-up to the story. It was vital for my purposes for participants to have quick and easy visual recognition of the characters in the stories. Therefore, I changed the originals, which had different names in each story, to use only two female and two male names. The order of introduction of the characters, and the character mentioned in sentences two and three, were counter-balanced across items, but were the same across conditions. There were 24 fillers, taken from a different experiment, also in Koornneef (2008). See Appendix for all stimuli. In the examples below, the regions of interest are italicized, and the critical manipulations are underlined.

(12) Luuk and Stijn volgen de cursus ‘Koken voor dommerdjes’. Vandaag <u>moesten zij een appeltaart bakken en hun projectjes</u> kan men het best omschrijven als ‘interessant’. Luuk durft *zijn* creatie zelfs niet te proeven, maar *Stijn wel*. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus, maar je weet maar nooit.

Luuk and Stijn follow the course ‘Cooking for dummies’. Today they had to bake an apple pie and their little projects can best be described as 'interesting'. Luuk dare not even taste *his* creation, but *Stijn does*. So far there are no deaths during the course, but you never know.

(13) Luuk and Stijn volgen de cursus ‘Koken voor dommerdjes’. Vandaag moest Luuk een appeltaart bakken en zijn projectje kan men het best omschrijven als ‘interessant’. Luuk durft *zijn* creatie zelfs niet te proeven, maar *Stijn wel*. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus, maar je weet maar nooit.

Luuk and Stijn follow the course ‘Cooking for dummies’. Today Luuk had to bake an apple pie and his little project can best be described as 'interesting'. Luuk dare not even taste *his* creation, but *Stijn does*. So far there are no deaths during the course, but you never know.

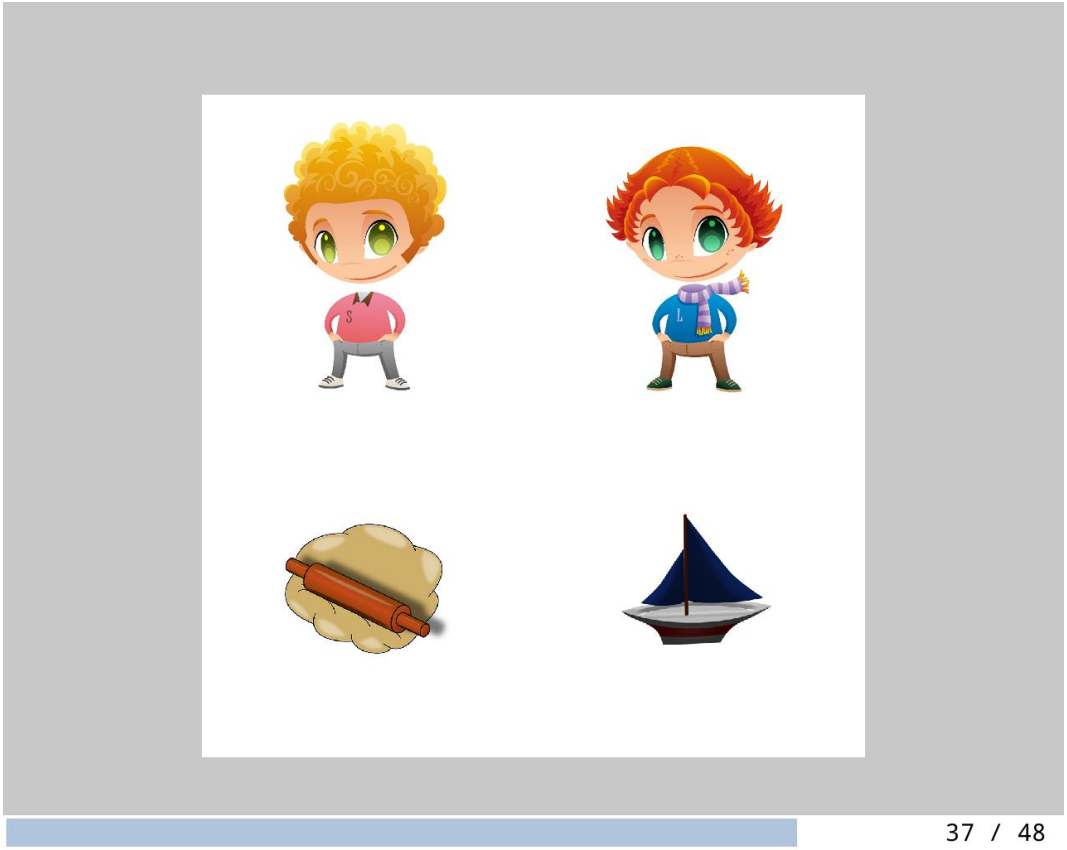
In the original experiment, the conditions were described as coreference biased and bound-variable biased, but it will be helpful here to add a layer of distinction between the two. The coreference biased condition is also, for our purposes, the low-ambiguity condition. In this condition, the one character is made the topic of the story in the critical manipulation. The bound-variable biased condition is also the high-ambiguity condition. Here, both characters are highlighted by the critical manipulation, so the non-antecedent may be more accessible.

Audio stimuli The stories were recorded as spoken by a male native speaker of Dutch, in a sound-proof booth, using version 2.0.0 of Audacity(R) recording and editing software. Another native speaker of Dutch listened in during the recording session, and sentences with speech errors or other disfluencies were re-recorded during the same session. Praat (Boersma 2001) was used to cut the long recordings into individual items, to remove noise, add silence at the beginning and end of each item, and to code the critical regions.

Visual stimuli consisted of cartoon-like drawings. The screen displayed 4 images during each trial. There were either two female characters or two male characters (permission for use kindly provided by Danilo Sanino), one related image, and one unrelated image. All images except for the four characters were licensed for free use. Images were all in color, and fit within 350 ^{x} 350 pixels, except for the four characters, which were slightly larger at 350 ^{x} 450 pixels. The images of the characters were altered only by adding the first letter of the character’s name onto their shirt, to aid participants in identification. The two characters appeared on the top half of the screen, with left/right placement counter-balanced across items but identical across conditions. The related and unrelated images

appeared on the bottom half of the screen, with left/right placement counter-balanced across items but identical across conditions.

Fig. 1. The image in this figure appeared with the example stimuli given above, and was identical in both conditions. The character on the top left is Stijn, identifiable by the "S" on his shirt, and the other is Luuk, identified by an "L". The image on the bottom left of the dough and rolling pin is the related image, as Stijn and Luuk are following the course "Cooking for Dummies" in the story. The sailboat is unrelated. Across the bottom of the image is a progress bar, showing the participant the total number of trials completed / total number of trials in the experiment.



Design. Each of the 24 experimental items occurred in two conditions. The resulting 48 items were divided into two lists, each containing 24 items (12 of each condition) and the 24 fillers, and participants were randomly assigned to a list.

Procedure. Participants were first administered a reading span task (Daneman & Carpenter 1980) as a background measure. They were given a book containing one sentence per page, with blank pages occurring after a certain number of sentences. The number of sentences between blank pages begins at three and increases to seven. They were instructed to read each sentence out loud, turn the page

immediately, read the next sentence, etc. until turning to a blank page. At the blank page, the task was to list the last word of every sentence, going back to the previous blank page, in the same order in which they were read. The test ended with a 'mercy rule' – once the participant failed to accurately remember 4 lists in a row, the test was stopped.

The particpants were then seated approximately 60 cm from a display monitor. The screen was set to a resolution of 1280 x 1024 pixels. Movements of the right eye were tracked with an SR Research Eyelink 1000 at 500 Hz (with the exception of one participant, whose left eye was used due to calibration issues). Calibration was performed before both the practice session and the experiment proper, using a 13-point calibration procedure. The practice session consisted of a brief introduction to the characters in the story followed by 10 practice items. Half of all items and fillers, including practice items, were followed by the question *Heb je dit woord gehoord?(Did you hear this word?)* followed by a word that either was or was not present in the just-heard story. The participant answered by pressing one of two buttons. A purple button on the left was a 'yes' response, and an orange button on the right was a 'no' response.The answers were analyzed to ensure all participants were attentive to the audio stimuli. Each trial began with a fixation point at the center of the screen. The image was presented for 200 milliseconds before the audio began. At the end of each trial, the image remained on the screen for 300 ms of silence before the next trial.

Auditory regions of interest. The audio stimuli were coded for five regions. In both conditions, the first region began at the first mention of the characters' names. In the coreference condition, the second region was the second mention of the character who was the target of the bias; in the bound-variable condition, the second region was a plural pronoun referring to both characters. The third region was a possessive pronoun; the fourth was the ellipsis site; and the fifth marked the end of the speech/beginning of the silence. The regions are shown in (14) and (15). The two main regions of interest are those labeled *pronoun* and *ellipsis*.

(14) *Bound-variable condition*

1st mention:	Luuk and Stijn volgen de cursus 'Koken voor dommerdjes'.
	Luuk and Stijn follow the course 'Cooking for dummies'.
plural:	zij een appeltaart bakken en hun projectjes
	they had to bake an apple pie and their little projects
pronoun:	*zijn* creatie zelfs niet te proeven, maar
	dare not even taste *his* creation, but
ellipsis :	*Stijn wel*. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus
	Stijn does. So far there are no deaths during the course

(15) *Coreference condition*

1st mention:	Luuk and Stijn volgen de cursus 'Koken voor dommerdjes'.
	Luuk and Stijn follow the course 'Cooking for dummies'.
2nd mention:	Luuk een appeltaart bakken en zijn projectje
	Luuk had to bake an apple pie and his little project
pronoun:	*zijn* creatie zelfs niet te proeven, maar
	dare not even taste *his* creation, but
ellipsis :	*Stijn wel*. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus
	Stijn does. So far there are no deaths during the course

It takes around 200 ms to plan and execute a saccade (Altmann & Kamide 2004), so the probability of fixating on an image due to linguistic input begins to increase 200 ms after onset of the input, at the earliest. For this reason, the regions were coded in Praat to begin 200 ms after actual auditory onset. For example, the pronoun region was coded to begin 200 ms after onset of the pronunciation of the pronoun.

Visual regions of interest. There were five possible regions in which a participant's gaze could fall: either of the two characters, the related object, the unrelated object, or the background. The background area was defined as anything outside of one of the four images. Fixations were assigned to visual regions of interest automatically by the program Fixation (Cozijn 2006). Drift was manually checked and adjusted.

Predictions. In the original experiment, one of the groups distinguished during the post-hoc analysis spent more time reading the overt pronoun, and less time reading the ellipsis. Under Koornneef's proposal – that his groups differed in whether or not they calculated both the bound-variable relationship and the coreference relationship, leading to the same result – the high-span group is expected to have an increased proportion of fixations on the contextually appropriate antecedent for a longer period of time than the low-span group. No significant increase for the contextually inappropriate antecedent is expected for either high- or low-span participants during the pronoun region. Predictions at the ellipsis region are the same as (ii), below.

Under the current hypothesis, the group who spent more time reading the overt pronoun did so because they were able to use more working memory resources to process the pronoun's ambiguity. The prediction this makes for VWP data is as follows: (i) the high-span participants are expected to show an increase in proportion of fixations on the contextually inappropriate antecedent in the pronoun region, even in the condition with low ambiguity; (ii) the high-span participants are

expected to show an advantage at the ellipsis site in the form of a higher proportion of fixations to the appropriate antecedent earlier than participants with low reading span scores.

Under the current hypothesis, there are two main possibilities for the low-span participants at the overt pronoun. Following Gernsbacher's (1989) finding that a pronoun automatically activates multiple possible antecedents, I expect the low-span participants, like the high span participants, to show an immediate increase in fixation proportion on the contextually inappropriate antecedent. Also, consistent with Nieuwland & Van Berkum's (2006) proposal that low-span readers may 'notice' the ambiguity, but not pursue the different possible interpretations due to a need to conserve processing resources, it is expected that fixation proportion for the inappropriate antecedent will rapidly decline. However, it is possible that the level of activation that the contextually inappropriate antecedent receives will not be enough to have an effect on looking behavior. This corresponds with Nieuwland & Van Berkum's (2006) alternative suggestion that low-span readers might not notice the ambiguity. In this case, the inappropriate antecedent is not expected to receive a significant increase in fixation proportion. During the ellipsis region, these same participants are expected to show a disadvantage in resolving the ambiguity of the unpronounced pronoun, in the form of a delay in increasing fixation proportions on the antecedent (as compared to the high-span group).

Pre-processing & analysis. Prior to analysis I conducted visual inspections of the data. The distractor images accounted for very few fixations overall. The related images attracted more attention, but the two characters by far accounted for the largest proportion of fixations. No single experimental item stood out as having a disproportionate number of looks to either the distractor or the related image.

I conducted the data processing and analysis with R version 3.1.2 (2014). The dependent variable was proportion of fixations on each of the two characters. The proportion of fixations on each character over time (in 2 millisecond time intervals) for each condition was calculated separately for each reading span group (i.e., high-span and low-span) by adding the number of trials in which each image was fixated at 2-ms intervals and dividing it by the total number of trials in which a fixation to any picture or location was observed during this time interval. Blinks and saccades were excluded. The two main auditory regions of interest were the pronoun and the ellipsis. For these two regions, proportional values for each image at each time interval were averaged across each group of participants, separately for each condition. To investigate within-subject changes in attention over time, I conducted planned Wilcoxon signed rank tests. This is a non-parametric test designed for testing two conditions when the same participants take part in each condition. A Wilcoxon rank sum test was used for between-group comparisons.

4.2.2 Results

Reading span. The traditional method of scoring reading span tests was used. This gives two scores – one is equal to the highest set size, and the other is equal to the total number of words in correct sets. Words that were correctly remembered but are contained in sets that were not remembered correctly were not counted. I here report only the scores from total words in correct sets because these were used in the statistical analysis. A median-split was used, yielding an approximately equal set of participants in each group. The high-span group consisted of participants with a reading span greater than 15; the low-span group consisted of participants with a reading span of 15 or less. Table 1 shows details for the reading span scores.

Table 1. Total number of words in correctly remembered sets.

Min	4
Max	65
Mean	18.4
Median	15
Mode	15
Sd	12.1

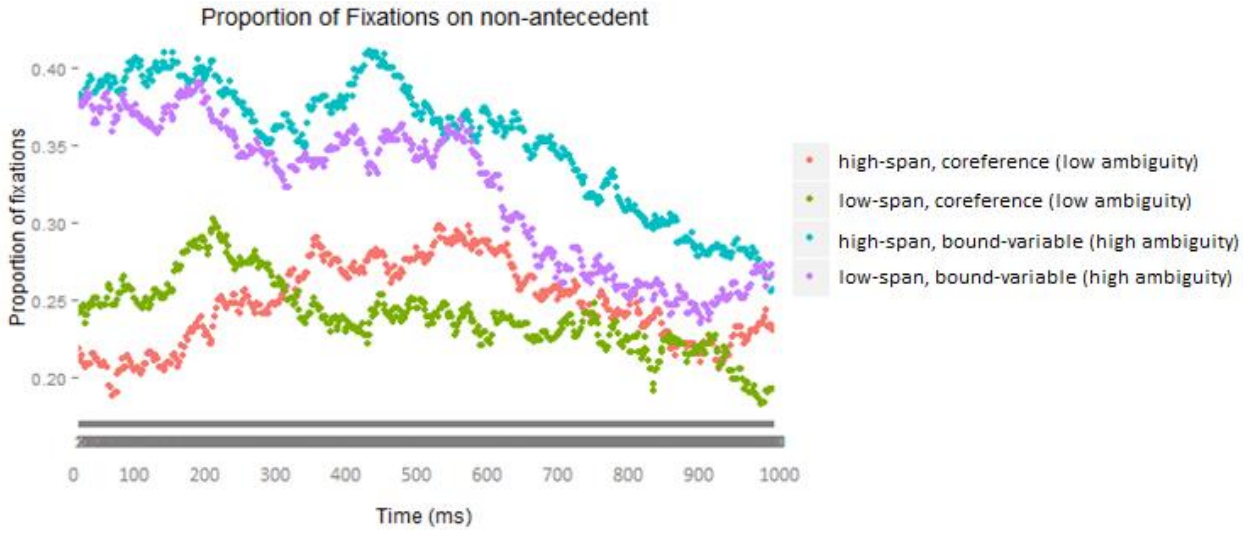
Probe words. Mean correct responses to probe words did not differ by item type, condition, or group. Table 2 contains shows percent correct.

Table 2. Percent correct responses to probe words

Min	71
Max	96
Mean	88

Overt pronoun. Figure 2 shows the proportion of fixations on the contextually inappropriate antecedent at the pronoun region. The main prediction for this region of interest is that the high-span group should show an increase in fixation proportion, and the low-span group may or may not show the same but, if they do, it will be short-lived.

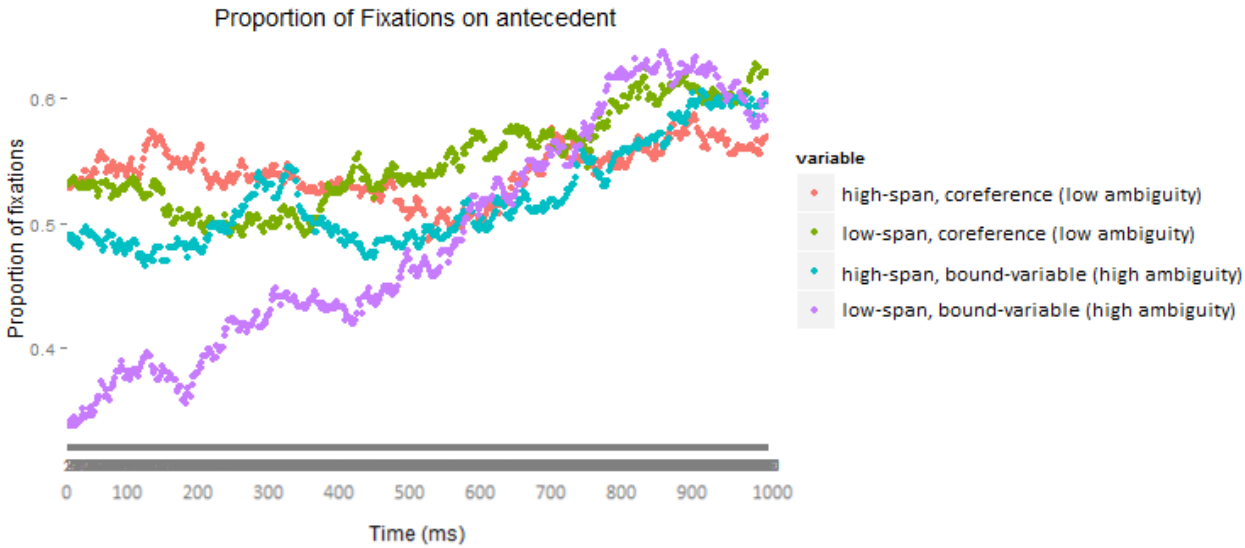
Fig. 2. Proportion of fixations on the contextually inappropriate antecedent while hearing the overt pronoun for both high and low reading span groups in each condition[8]*. Each combination of group and condition is marked by color.*



There is a clear difference between the two conditions. This data suggests that both high- and low-span groups attend more to the non-antecedent in the condition with higher ambiguity. What is of interest, though, is the difference between the high- and low-span participants in the coreference (low ambiguity) condition (green and salmon colored lines). Visual inspection of the graph shows a pattern consistent with the non-antecedent being activated for both groups upon hearing the pronoun. According to a Wilcoxon signed rank test, the increase between 100 and 200 milliseconds post-pronoun-onset is significant for high-span participants ($p < .001$) with a large effect size ($r = -0.76$) and for the low-span participants ($p < .001$), also with a large effect size ($r = -0.87$). For the low-span group, this activation is very brief, significantly decreasing by 300 milliseconds ($p < .001$, $r = -0.84$). The high-span group maintains the higher proportion of looks for several hundred milliseconds, with a significant decrease in fixation proportion occurring by 700 ms ($p < .001$, $r = -0.87$). This can be interpreted as being in line with the prediction that high-span readers would spend resources processing the ambiguity, and the low-span readers would either not notice it, or not process it. The data appears to indicate that they do not process it, but that a boost in activation does occur, similar to Gernsbacher's (1989) finding that a pronoun immediately activates multiple possible antecedents, even in a disambiguating context. In order to determine what effect, if any, this has on activation of the contextually appropriate antecedent, I also analyzed the data for the corresponding character, as shown in Figure 3.

[8] For all graphs illustrating the data in the pronoun region, pronoun onset begins 200 ms prior to zero on the x-axis (see *Auditory regions of interest*, section 4.2.1); pronoun offset occurs between 100-200 ms in all auditory stimuli. The initial changes in fixation proportions that are directly related to processing the pronoun should begin between 0 and 300 ms (Altmaan & Kamide 2004).

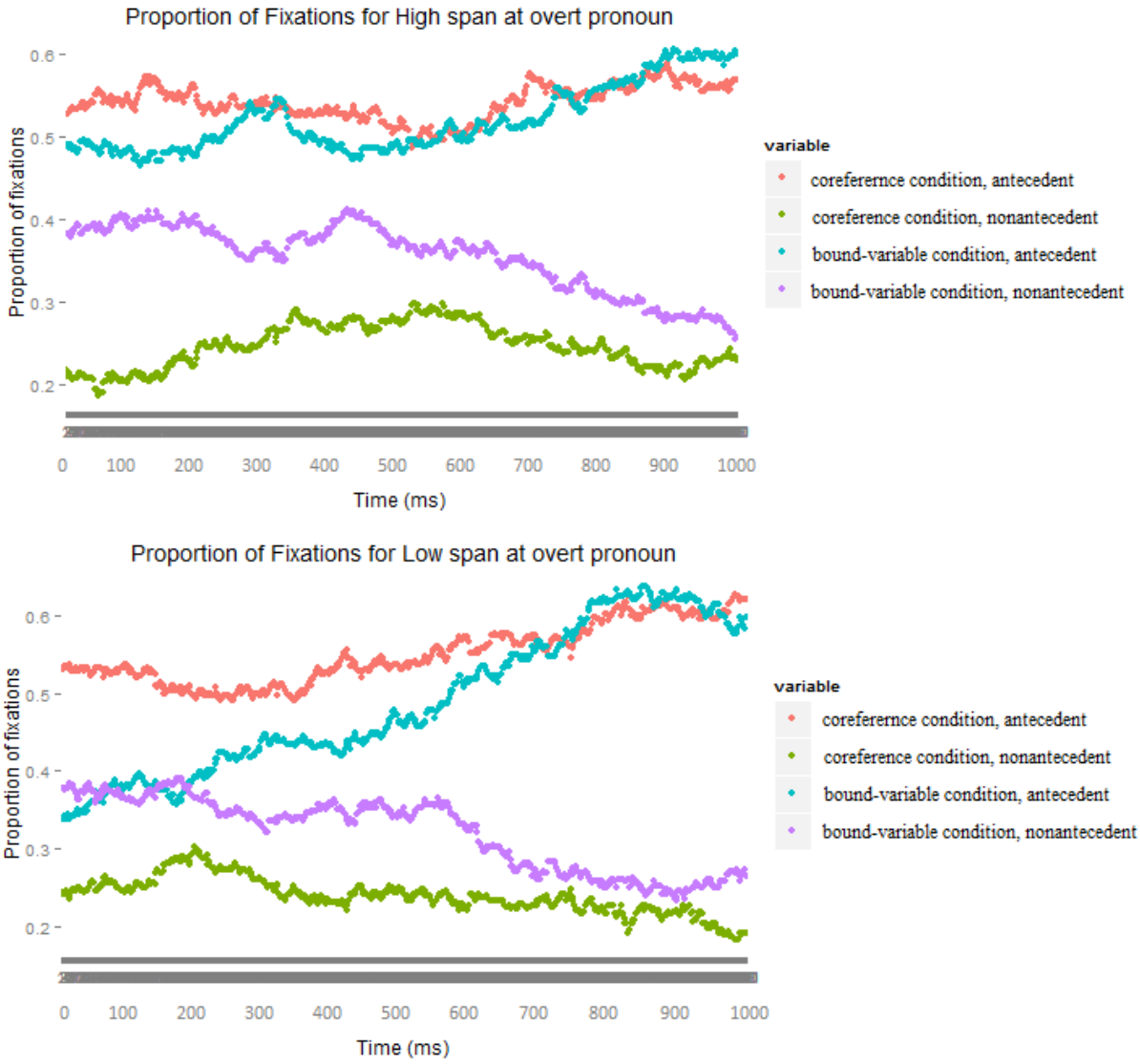
Fig. 3. Proportion of fixations on the contextually appropriate antecedent while hearing the overt pronoun for both high and low reading span groups in each condition. Each combination of group and condition is marked by color.



A Wilcoxon signed-rank test on within participant effects in the first 200 ms reached significance ($p < .001$) with a large effect size ($r = -0.87$) for the low-span group between the two conditions (the green and purple lines in Fig. 3). The high-span group does not show significant differences between the two conditions. A Wilcoxon rank-sum test between the high- and low-span groups in the bound-variable condition (the blue and purple lines in Fig. 3) were not significant in the first 200 ms.

The two graphs in Figure 4 below show an interesting difference between the high- and low-span groups in the bound-variable/high ambiguity condition (the blue and purple lines).

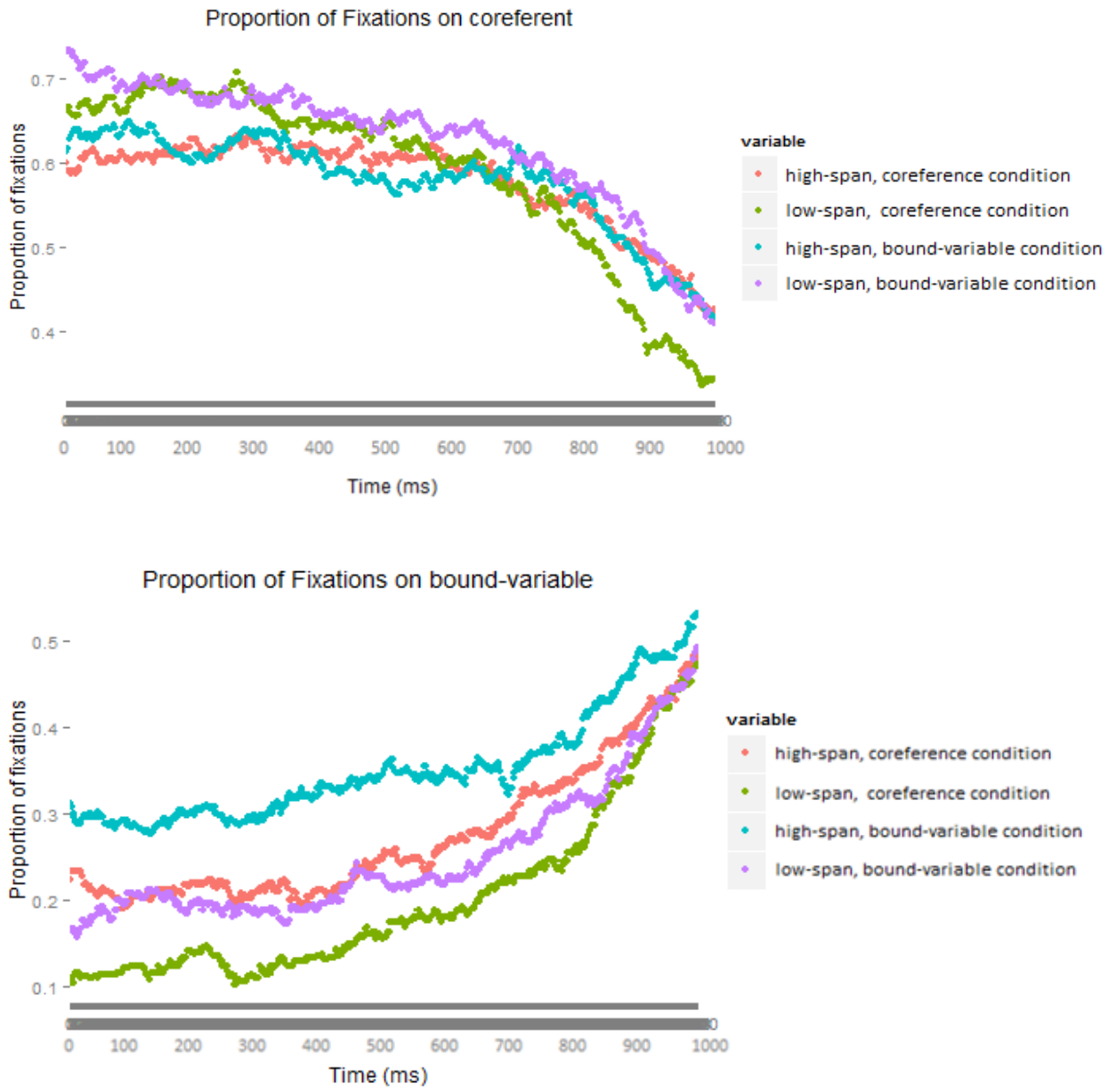
Fig. 4. The upper graph shows the proportion of fixations for the high-span group at the overt pronoun. The lower graph is for the low-span group. Like the graphs above, condition is marked by color; unlike the graphs above, antecedent vs. non-antecedent is also marked by color.



In the bound-variable (high-ambiguity) condition, the low-span group has approximately equal fixation proportions on both possible antecedents (i.e., no statistically significant difference) at the beginning of the pronoun region, while the high-span group shows more attention to the contextually appropriate antecedent relative to the inappropriate antecedent ($p < .001$, $r = -0.87$).

Ellipsis site. Visual inspection of the graph in Figure 5 shows that there appears to be a strong bias for the coreferential interpretation in both conditions and across reading span scores.

Fig. 5. Proportion of fixations on the character representing the coreferential interpretation while hearing the ellipsis. For example, looks to Luuk while hearing "Stijn wel."[9] *Line color indicates the combination of reading span group and condition.*



In the bound-variable condition, proportion of fixations on the character representing the coreferential interpretation is above .60, for both low- and high-span participants. Proportion of fixations to the other character is around .20 for low-span and .30 for the high-span group. Wilcoxon tests at 200 and 500 ms were $p < .001$, with similar effect sizes, each rounding to $r = -0.87$. It is especially unexpected that participants would fixate on the character who is not the subject of bias in the bound-variable

[9] For all graphs illustrating the data in the ellipsis region, ellipsis onset begins 200 ms prior to zero on the x-axis (see *Auditory regions of interest*, section 4.2.1).

condition for two reasons. The first is that the bound-variable interpretation is the less costly of the two according to the economy hierarchy. As such, it is expected always to be computed first, and then repaired when it is discovered not to be in line with information from the discourse. Based on this, the exact opposite of the obtained pattern of results might be anticipated (people would show a general preference for the bound-variable interpretation of the elided pronoun regardless of condition). The second, more straightforward, reason is that the character who is the subject of bias in the bound-variable condition is the same character whose name is heard at the beginning of the ellipsis region. For example, when the participant hears *Stijn wel* in the bound-variable condition, Stijn is the antecedent of the elided pronoun. It is striking that this character is not attended to because the VWP is based on the tendency for people to look at images of or related to what they are processing from linguistic input. So, given that the character's name was spoken overtly *and* the character is the antecedent that the condition biases toward, it was expected that there would be a strong pattern of attention to that character. This indicates a possible confound, which could also be the reason for the findings at the pronoun region, rather than ambiguity. This will be discussed more fully in the Discussion below.

4.3 Discussion

Reuland's (2001) Principles of Binding (PoB) framework proposes that there are multiple routes to comprehending anaphora. Its economy hierarchy predicts the relative cognitive cost of each route according to the amount of working memory resources required for successful anaphor resolution. Koornneef (2008) conducted a set of experiments designed to determine whether the predictions made by the PoB are instantiated in psychological reality. The results of an eye-tracking reading experiment offered support for the PoB framework and economy hierarchy. However, in a post-hoc analysis, Koornneef found an effect of individual differences that seemed antithetical to the principle of economy. Based on reading times, some participants appeared to voluntarily calculate the more costly route. However, pronouns are inherently ambiguous, and some studies suggest that availability of working memory resources is a factor in resolving ambiguity. I proposed that Koornneef's (2008) post-hoc findings may be rooted in working memory capacity, and replicated his eye-tracking experiment to investigate this proposal. I used the visual world paradigm instead of the reading paradigm in order to determine whether looking behavior can track pronoun resolution processes in real time, potentially leading to the ability to detect the role of working memory capacity in pronoun resolution. I predicted that participants with high reading span scores would (i) show an increase in proportion of fixations, even in the bound-variable condition, (ii) would show an earlier increase in proportion of fixations on the appropriate antecedent during the ellipsis region, relative to

participants with low reading span scores. I predicted that the low-span group would (i) either have no increase in fixation proportion on the contextually inappropriate antecedent in the pronoun region (i.e., not notice the ambiguity), or, show only a fleeting increase (i.e., not process the ambiguity); (ii) take longer than the high-span group to fixate on the appropriate antecedent in the ellipsis region.

In the pronoun region, the results of the current experiment suggest a behavioral pattern that is in accordance with Nieuwland & Van Berkum's (2006) proposal that people with high reading span, indicative of high working memory resources, are more attentive to ambiguity, and that this is true even for so-called unambiguous pronouns. In the present study, the high-span group of participants showed a sustained boost in attention to (fixations on) the contextually inappropriate antecedent. The low-span group showed a similar, but briefer, attention boost, suggesting that both antecedents are accessed, perhaps automatically, but not processed in order to conserve resources. If the longer reading times at the pronoun region in Koornneef's (2008) data had been the result of some participants spending extra processing resources on calculating both routes to the same antecedent, then there should not be an increase in fixations on the contextually inappropriate antecedent; rather there should be a sustained increase in proportion of fixations to the appropriate antecedent in the high-span group compared to the low-span group in the bound-variable condition, which was not the case. However, there was a confound, as mentioned in section 4.2.2, which I return to now. The evidence for the confound was in the results for the ellipsis region.

Phillips & Parker wrote "we think that evidence on what information is accessed at an ellipsis site does not indicate what representations are constructed at an ellipsis site" (2013:16). The current study suggests that they are right. The expectations for looking behavior were not met at the ellipsis site, despite being the same for both Koornneef's (2008) and the current hypotheses. This is concerning because the visual world paradigm is predicated on the phenomenon that eye gaze is typically directed at images of or related to what is being processed from linguistic input. If participants see an image of a boat and a car, and they hear the word *sail* they will look to the boat with far greater frequency. In the present study, this expectation was not fulfilled in the ellipsis region. This region began with the name of the character representing the bound-variable reading of the elided pronoun, followed by the ellipsis site itself which, in the bound-variable condition, should also elicit looks to the same character. However, across both conditions, there was a strong tendency to fixate on the character representing the coreference reading. One possibility is that the intonation of the speaker may be more consistent with a coreferential treatment of the elided pronoun. An informal test suggests this is not the case, but the question should be tested more formally.

A stronger possibility is a confound of topichood. The *centering framework* developed by Grosz, Joshi & Weinstein (1995) was specifically intended to model *attentional state*, meaning the focus of attention of the discourse participants at any point during the discourse. The authors identified two types of focus – local and global. Local focus occurs at the sentence or clause level,

global focus concerns the wider discourse. The linguistic stimuli used in the current study may have a global focus that takes one character as its *center*[10], even in the stories that are meant to be about both characters. Essentially, this one character is the topic of the discourse. This may create enough command of the attentional state to affect looking behavior. In the bound-variable/high ambiguity example stimulus in (16), each referential expression (either a name or pronoun) occurring before the ellipsis-containing clause is underlined.

(16) Luuk and Stijn volgen de cursus 'Koken voor dommerdjes'. Vandaag moesten zij een appeltaart bakken en hun projectjes kan men het best omschrijven als 'interessant'. Luuk durft zijn creatie zelfs niet te proeven, maar Stijn wel. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus, maar je weet maar nooit.

Luuk and Stijn follow the course 'Cooking for dummies'. Today they had to bake an apple pie and their little projects can best be described as 'interesting'. Luuk dare not even taste his creation, but Stijn does. So far there are no deaths during the course, but you never know.

The character representing the coreferential interpretation of the elided pronoun (here, *Luuk*) is referred to by name twice, and by pronoun three times. The character representing the bound-variable interpretation of the elided pronoun (here, *Stijn*) is referred to by name once and by pronoun twice, but note that both pronouns referring to this character are plural. He is not referred to singly until the ellipsis. By the time the ellipsis is heard, Luuk has a greater frequency of being a target of reference, and is furthermore the topic of the sentence that introduces the ellipsis in its second clause. Some corroboration for the topichood hypothesis comes from a recent VWP study by Colonna, Schimke & Hemforth (2014). They looked at whether and how topicalization[11] affects ambiguous pronoun resolution, with audio stimuli exemplified in (17).

(17) Here are the postman and the pirate.
a. As for the postman$_i$, he$_i$ will slap/slapped the pirate when he will be/was home.
b. As for the postman$_i$, the pirate will slap/slapped him when he$_i$ will be/was home.

[Colonna et al 2014:181]

The authors concluded that "topicalization, if it is clearly marked, thus seems to make a potential antecedent highly accessible for an ambiguous pronoun, at least across the languages we investigated

[10] The authors describe the term *center* as a strictly semantic object, not a word or a syntactic form. (Grosz, Joshi & Weinstein (1995).
[11] The authors indicate that their intended meaning of *topicalization* is the same as in Grosz's work. See Colonna, Schimke & Hemforth (2014:192).

so far" (p. 191). The languages in question are French and German. Given the relation of Dutch to German, it seems reasonable to suggest Dutch may function similarly, but this remains to be tested. Colonna et al (2014) also tracked thr preference through time and found that increased probability of fixating on the topicalized character lasted until approximately 800 ms after pronoun onset when that character's grammatical role was subject; even longer when the topicalized character was an object. That's not to suggest that in the current study the pronoun was necessarily resolved in favor of the coreferential interpretation even in the bound-variable condition, but its topichood status made it highly activated in working memory. As such, it is likely that attention to the contextually inappropriate antecedent during the ellipsis region reflects discourse-level processing, which may take precedence in the command of visual attention. Also note in (16) that the pronoun region coincides with the last of five expressions referring to the same character,[12] thus confounding ambiguity with topichood at the pronoun region as well and making any interpretation of the results in terms of ambiguity impossible at this time. Future studies may be able to make use of the centering framework to construct linguistic stimuli that manipulate topichood separately from ambiguity to look at the question of whether pronouns automatically activate multiple, contextually-possible antecedents, as Gernsbacher's (1989) and Nieuwland & Van Berkum's (2006) work suggests, or whether their results may also be an effect of attentional state related to topichood, rather than activation by the pronoun itself.

A final note: the visual world paradigm may not be an effective tool for investigating finer-grained linguistic processes within a discourse context, such as computing a bound-variable reading of a pronoun or resolving ambiguity, because they can become obscured by other simultaneous linguistic processes. The VWP is meant to look at the relationship between comprehension and visual attention. That people typically look at images related to what they are processing does not mean that they always do, especially if higher- and lower-order processes produce conflicting demands on attention. The higher-order processes, like discourse comprehension, may generally win out over lower-order processes, like pronoun resolution. More studies along the lines of Colonna et al (2014) may help to disentangle this confound by carefully constructing linguistic stimuli to vary the amount of bias introduced to the pronoun resolution process by topichood versus the economy hierarchy.

[12] In whole in the case of the names and singular pronouns; in part in the case of the plural pronouns: *Luuk…they…their…Luuk…his*.

References

Audacity® software is copyright 1999-2015. Available at http://audacity.sourceforge.net/. It is free software distributed under the terms of the GNU General Public License. The name Audacity® is a registered trademark of Dominic Mazzoni.

Altmann, G.T.M. & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision and action: Eye movements and the visual world* (347-386). New York: Psychology Press.

Arnold, J. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, *31*(2), 137–162.

Arnold, J., Eisenband, J.G., Brown-Schmidt, S. & Trueswell, J.C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eye tracking. *Cognition*, *76*, B13-B26.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*, 341-345.

Clark, H.H. & Sengul, C.J. (1979). In search of referents for nouns and pronouns. *Memory and Cognition*, *7*, 35-41.

Colonna, S., Schimke, S. & Hemforth, B. (2014). Information structure and pronoun resolution in German and French: Evidence from the visual-world paradigm. In *Psycholinguistic approaches to meaning and understanding across languages* (pp. 175-195). Springer International Publishing.

Cooper R.M. (1974). The control of eye fixation by the meaning of spoken language: a new methodology for the real-time investigation of speech perception, memory and language processing. *Cognitive Psychology*, *6*, 84–107.

Cozijn, R. (2006). Het gebruik van oogbewegingen in leesonderzoek [The use of eye movements in reading research]. Tijdschrift voor Taalbeheersing, 28, 220–232.

Cozijn, R., Commandeur, E., Vonk, W., Noordman, L.G.M. (2011). The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language, 64*, 381-403.

Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

Fiengo, R. & May, R. (1994). *Indices and identity*. MIT press.

Frazier, L. & Clifton Jr, C. (2000). On bound variable interpretations: The LF-only hypothesis. *Journal of Psycholinguistic Research, 29*(2), 125-140.

Frazier, L. & Clifton Jr, C. (2001). Parsing coordinates and ellipsis: Copy α. *Syntax, 4*(1), 1-22.

Friederici, A.D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences, 6*(2), 78-84.

Friederici, A.D., Steinhauer, K. & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition,27*(3), 438-453.

Gernsbacher, M.A. (1989). Mechanisms that improve referential access. *Cognition, 32*, 99-156.

Grosz, B.J., Weinstein, S. & Joshi, A.K. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational linguistics, 21*(2), 203-225.

Hahne, A. & Friederici, A.D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research, 13*. (2002) 339-356.

Hardt, D. (1993). *Verb phrase ellipsis: Form, meaning and processing* (Doctoral dissertation). University of Pennsylvania.

Heim, I. (1998). Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. *MIT working papers in linguistics, 25*, 205-246.

Huettig, F., Rommers, J. & Meyer, A.S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151-171.

Kaiser, E. (2003). *The quest for a referent: A crosslinguistic look at reference resolution*. (Ph.D. dissertation). University of Pennsylvania, Philadelphia, PA.

Kaiser, E. (2005). When salience isn't enough: Pronouns, demonstratives and the quest for an antecedent. In Laury, R., (Ed.). Minimal reference in Finnic: The use and interpretation of pronouns and zero in Finnish and Estonian discourse. Helsinki, Finland: Suomalaisen Kirjallisuuden Seura, 135-162.

Kaiser, E., Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition, 112*(1), 55-80.

Kaiser, E., & Trueswell, J. (2004). The referential properties of Dutch pronouns and demonstratives: Is salience enough. *Arbeitspapier Nr.117*, 137.

Kaiser, E.; Runner, J.T. (2008). Intensifiers in German and Dutch anaphor resolution. In: Abner, N., Bishop, J., (Eds.), *Proceedings of the 27th West Coast Conference on Formal Linguistics*. (265-273). Somerville, MA: Cascadilla Proceedings Project.

Kennedy, C. (2003). Ellipsis and syntactic representation. *The Interfaces: Deriving and Interpreting Omitted Structures*, 29-53.

Kennedy, C. & Lidz, J. (2001). A (covert) long distance anaphor in English. Paper presented at the *Proceedings of WCCFL, 20* 318-331.

Koornneef, A.W. (2008). Eye-catching anaphora. (PhD dissertation). LOT.

Koornneef, A.W. (2010). Looking at anaphora: The psychological reality of the primitives of binding model. In: M. Everaert, T. Lentz, H. de Mulder, Ø. Nilsen & A. Zondervan (Eds.), *The linguistics enterprise: From knowledge of language to knowledge in linguistics* (pp. 141–166). Amsterdam: John Benjamins.

Koornneef, A.W., Avrutin, S., Wijnen, F. & Reuland, E. (2011). Tracking the preference for bound-variable dependencies in ambiguous ellipses and only. *Syntax and Semantics, 37*, 67-100.

Koornneef, A.W. & Sanders, T.J. (2013). Establishing coherence relations in discourse: The influence of implicit causality and connectives on pronoun resolution. *Language and Cognitive Processes, 28*(8), 1169-1206.

Koornneef, A.W. & Van Berkum, J.J.A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye-tracking. *Journal of Memory and Language*, *54*, 445–465.

Martin, A.E. & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language, 58*(3), 879-906.

MacDonald, J.L. & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, *34*, 543-566.

Murphy, G.L. (1985). Psychological explanations of deep and surface anaphora. *Journal of Pragmatics, 9*(6), 785-813.

Nickels, S., Bokhari, F.S. & Steinhauer, K. (2014). 'Semantic blocking' solved at last: ERP evidence challenging syntax-first models. (http://bit.ly/1Ig0VDK)

Nieuwland, M.S. (2014). "Who's he?" Event-related brain potentials and unbound pronouns. *Journal of Memory and Language, 76*, 1-28.

Nieuwland, M.S. & Van Berkum, J.J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098-1111.

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language*, *59*, 494–522.

Phillips, C. & Parker, D. (2014). The psycholinguistics of ellipsis. *Lingua*, *151*, 78-95.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reinhart, T. (1983). *Anaphora and Semantic Interpretation*. London: Croom Helm.

Reinhart, T. (1991). Elliptic conjunctions: Non-quantificational LF. In A. Kasher (Ed.,) *The Chomskyan Turn* (pp. 360-384). Blackwell, Oxford.

Reinhart, T. (2000). Strategies of anaphora resolution. In H. Bennis, M. Everaert and E. Reuland (Eds.). *Interface Strategies* (pp. 295-325). Amsterdam: Royal Academy of Sciences.

Reinhart, T. & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, *24*, 657-720.

Reuland, E. (2001). Primitives of Binding. *Linguistic Inquiry*, *32*, 439-492.

Sanino, Danilo. ddraw illustrations. http://ddraw.blogspot.it/.

Shapiro, L.P., Hestvik, A., Lesan, L. & Garcia, A.R. (2003). Charting the time-course of VP-ellipsis sentence comprehension: Evidence for an initial and independent structural analysis. *Journal of Memory and Language, 49*(1), 1-19.

St. George, M., Mannes, S. & Hoffman, J. (1997). Individual differences in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience 9*(6), 776-787.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior,* 18, 645-659.

Swinney, D., Prather, P. & Love, T. (2000). Time-course of lexical access and the role of context: Converging evidence from normal and aphasic processing. *Language and the brain: Representation and processing*, 273-292.

Tanner, D. (2013). Individual differences and streams of processing. *Linguistic Approaches to Bilingualism, 3*(3), 350-356.

Tanner, D. (2014). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement.

Tanner, D. & Van Hell, J.G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*, 289-301.

Vasić, N. (2006). Pronoun comprehension in agrammatic aphasia: The structure and use of linguistic knowledge. (PhD dissertation). LOT.

Vasić, N., Avrutin, S. & Ruigendijk, E. (2006). Interpretation of pronouns in VP-ellipsis constructions in Dutch Broca's and Wernicke's aphasia. *Brain and Language, 96*(2), 191-206.

Appendix

Coreference Condition

1. Roos en Anouk proberen een vriendje te vinden door afspraakjes te maken via het internet. Roos heft sinds kort een verhouding met zo'n internetdate en ziet alles zeer rooskleurig in. Roos ziet toekomst in haar relatie, maar Anouk helaas niet. Liefde blijft natuurlijk altijd een gok.

2. Anouk en Roos houden wel van een house-feestje. Anouk neemt op de grote feesten altijd een XTC-pilletje. Van de twee gabbertjes vindt Anouk dat haar gedrag dan bijna niet verandert, maar Roos wel. Sommigen hebben het blijkbaar allemaal niet zo door.

3. Roos en Anouk zijn op een ontzettend saaie verjaardag in Lunteren. Roos kent niemand en probeert het ijs te breken door grapjes te maken, hetgeen zij met veel zelfvertrouwen doet. Van de twee heeft Roos het idee dat haar humor ook daadwerkelijk gewaardeerd wordt, maar Anouk niet. Het valt te hopen dat het niet zal eindigen in een avond vol schaamte en ergernissen.

4. Sinds lange tijd gaan Anouk en Roos elke week wel een keer naar de sauna. Anouk heeft in de loop der jaren een enorm dikke buik gekregen. Van de twee heeft Anouk geen problemen met haar buik, maar Roos wel. Ik dacht begrepen te hebben dat je ontzettend afviel in de sauna, niet altijd blijkbaar...

5. Luuk en Stijn zijn gezegend met een berg acteertalent. Luuk wordt daarom vaak gecast voor aparte rollen, al verschilt hij daarover van mening. Luuk vindt zijn rollen namelijk altijd normaal, maar Stijn niet echt. 'Normaal' valt nu eenmaal niet te definieren.

6. Stijn en Luuk zijn lekker aan het borrelen. Stijn kletst een beetje over zijn werk als accountant en hij vindt het blijkbaar superleuk. Het probleem is dat Stijn zijn beroep interessant vindt, maar Luuk niet. Ze besluiten maar snel ergens anders over te gaan praten.

7. Roos en Anouk zijn dol op de zender MTV. Roos kon haar geluk niet op toen zij mee mocht doen aan het programma ‘Pimp My Room’, waarin haar kamer werd ‘opgeknapt’. Maar helaas, van de twee vriendinnen vindt Roos dat haar ‘gepimpte’ kamer klasse heeft, maar Anouk niet. Smaken verschillen nu eenmaal.

8. De studenten Stijn en Luuk zijn sinds kort kamergenootjes. Stijn is echt een TV-junkie en brengt hele dagen op de bank door. Van de twee denkt Stijn dat het zijn creativiteit stimuleert, maar Luuk niet. TV kijken kan blijkbaar echt een verslaving worden.

9. Luuk en Stijn zijn al lang collega's en goede vrienden. Zoals iedereen weet, heeft Luuk wel eens een nieuwe radiozender opgericht. Luuk wist zeker dat zijn nieuwe zender een succes zou worden, maar Stijn niet. Wat een zelfvertrouwen hebben sommigen mensen toch.

10. Luuk en Stijn komen elkaar tegen in het park. Luuk heeft zoals altijd een enorme hond aan de lijn. Luuk kan zijn hond onder controle houden, maar Stijn niet. Honden luisteren nu eenmaal niet naar iedereen.

11. Stijn en Luuk keren steeds weer terug in mijn verhaaltjes. Stijn heeft altijd zijn mond vol over ‘de arbeider’ en ‘de natuur’. Stijn heeft de overtuiging dat zijn ideeën echt socialistisch zijn, maar Luuk niet. Het is tegenwoordig de vraag of het echte socialisme nog wel bestaat.

12. Luuk en Stijn volgen de cursus ‘Koken voor dommerdjes’. Vandaag moest Luuk een appeltaart bakken en zijn projectje kan men het best omschrijven als ‘interessant’. Luuk durft zijn creatie zelfs niet te proeven, maar Stijn wel. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus, maar je weet maar nooit.

13. Anouk en Roos gaan elk jaar snowboarden in Aspen en zijn er dol op. Anouk maakte op de laatste skivakantie een sprong over een minuscuul schansje. Anouk vond haar sprong stoer, maar Roos niet. Echt hoog springen, dat kunnen de locals die het hele jaar niets anders doen.

14. Stijn en Luuk kennen elkaars grootouders vrij goed. Stijn heeft een beetje een gekke opa en daar denken ze zeer verschillend over. Stijn kan zijn vreemde opa niet waarderen, maar Luuk wel. Waarschijnlijk is het gewoon een vorm van schaamte.

15. Roos en Anouk hebben sinds kort het chatten op het internet ontdekt. Roos slaapt sindsdien nog maar maximaal twee uur per dag en maakt zich daar een beetje zorgen over. Roos denkt dat haar chat-gedrag niet normaal is, maar Anouk gek genoeg wel. Het valt te hopen dat ze het allebei een beetje binnen de perken kunnen houden.

16. Anouk en Roos zijn allebei ontzettend dol op koffie. Anouk maakte afgelopen zondag voor het eerst een koffie verkeerd, naar haar mening met veel 36uccess. Van het tweetal vond Anouk haar probeerseltje lekker, maar Roos niet. Misschien kun je in sommige gevallen echt spreken van een koffie verkeerd.

17. Roos en Anouk hebben in de loop der tijd veel gouden sieraden verzameld. Grappig genoeg heeft Roos op een bepaald moment in haar leven twee gouden tanden gehad. Roos vond haar gouden tanden ook daadwerkelijk mooi, maar Anouk niet. Gelukkig wordt iedereen op een gegeven moment volwassen.

18. Anouk en Roos spelen samen in een punkbandje. Anouk bespeelt een normale gitaar en Roos de basgitaar. Anouk ging laatst helemaal los tijdens haar solo. Van beide 'gitaarvirtuozen' was Anouk niet tevreden over haar timing en algemene performance tijdens de solo, maar Roos wel. De ene dag gaat het nu eenmaal voor je gevoel wat beter dan de andere.

19. Luuk en Stijn zitten samen op de filmacademie. Luuk heeft vandaag de laatste hand gelegd aan zijn nieuwste korte film. Van de twee jonge cineasten vindt Luuk dat zijn film goed gelukt is, maar Stijn niet. Uiteindelijk gaat het er natuurlijk wel om wat het publiek ervan vindt.

20. Stijn en Luuk kwamen toevallig tegelijk terug van een lange reis. Stijn gaf bij terugkomst een feestje voor zijn vrienden, maar was niet echt tevreden met de avond. Van hen beiden vond Stijn zijn feestje niet geslaagd, maar Luuk wel. Soms verheug je je gewoon teveel op iets.

21. Toen ze samen aan het reizen waren door Australië, hebben Roos en Anouk leren surfen. Roos wankelde voortdurend op het surfboard. Roos dacht dat haar surfstijl professioneel aandeed, maar Anouk niet. 'Hang loose' en 'no worries' zijn veelgehoorde kreten op de Australische stranden, maar niet altijd van toepassing op onhandige West-Europese meisjes.

22. Luuk en Stijn zijn net jarig geweest. Luuk heeft een fiets gekregen en is er erg blij mee. Luuk vindt zijn nieuwe fiets mooi, maar Stijn helaas niet. Toch hebben ze afgesproken morgen een leuk tochtje te maken.

23. Anouk en Roos waren allebei op zoek naar een nieuwe computer. Anouk heeft zich laten verleiden door een aanbieding bij de Mediamarkt en heeft een laptop aangeschaft. Anouk vindt haar laptop handig, maar Roos absoluut niet. Soms is het beter een wat duurder product te kopen.

24. Anouk en Roos waren allebei op zoek naar een nieuwe computer. Zij hebben zich laten verleiden door een aanbieding bij de Mediamarkt en hebben een laptop aangeschaft. Anouk vindt haar laptop handig, maar Roos absoluut niet. Soms is het beter een wat duurder product te kopen.

Bound-variable Condition

1. Roos en Anouk proberen een vriendje te vinden door afspraakjes te maken via het internet. Zij hebben allebei sinds kort een verhouding met zo'n internetdate, maar zien het niet even rooskleurig in. Roos ziet toekomst in haar relatie, maar Anouk helaas niet. Liefde blijft natuurlijk altijd een gok.

2. Anouk en Roos houden wel van een house-feestje. Zij nemen op de grote feesten altijd een XTC-pilletje. Van de twee gabbertjes vindt Anouk dat haar gedrag dan bijna niet verandert, maar Roos wel. Sommigen hebben het blijkbaar allemaal niet zo door.

3. Roos en Anouk zijn op een ontzettend saaie verjaardag in Lunteren. Zij kennen niemand en proberen allebei het ijs te breken door grapjes te maken, hetgeen zij niet met evenveel zelfvertrouwen doen. Van de twee heeft Roos het idee dat haar humor ook daadwerkelijk gewaardeerd wordt, maar Anouk niet. Het valt te hopen dat het niet zal eindigen in een avond vol schaamte en ergernissen.

4. Sinds lange tijd gaan Anouk en Roos elke week wel een keer naar de sauna. Zij hebben allebei in de loop der jaren een enorm dikke buik gekregen. Van de twee heeft Anouk geen problemen met haar buik, maar Roos wel. Ik dacht begrepen te hebben dat je ontzettend afviel in de sauna, niet altijd blijkbaar...

5. Luuk en Stijn zijn gezegend met een berg acteertalent. Zij worden daarom vaak gecast voor aparte rollen, al verschillen zij daarover van mening. Luuk vindt zijn rollen namelijk altijd normaal, maar Stijn niet echt. 'Normaal' valt nu eenmaal niet te definieren.

6. Stijn en Luuk zijn lekker aan het borrelen. Zij kletsen een beetje over hun werk als accountant, maar zij vinden het blijkbaar niet even leuk. Het probleem is dat Stijn zijn beroep interessant vindt, maar Luuk niet. Ze besluiten maar snel ergens anders over te gaan praten.

7. Roos en Anouk zijn dol op de zender MTV. Zij konden hun geluk niet op toen zij mee mochten doen aan het programma ‘Pimp My Room’, waarin hun twee kamers werden ‘opgeknapt’. Maar helaas, van de twee vriendinnen vindt Roos dat haar ‘gepimpte’ kamer klasse heeft, maar Anouk niet. Smaken verschillen nu eenmaal.

8. De studenten Stijn en Luuk zijn sinds kort kamergenootjes. Zij zijn echt TV-junkies en brengen hele dagen op de bank door. Van de twee denkt Stijn dat het zijn creativiteit stimuleert, maar Luuk niet. TV kijken kan blijkbaar echt een verslaving worden.

9. Luuk en Stijn zijn al lang collega's en goede vrienden. Zoals iedereen weet, hebben zij allebei wel eens een nieuwe radiozender opgericht. Luuk wist zeker dat zijn nieuwe zender een succes zou worden, maar Stijn niet. Wat een zelfvertrouwen hebben sommigen mensen toch.

10. Luuk en Stijn komen elkaar tegen in het park. Zij hebben zoals altijd allebei een enorme hond aan de lijn. Luuk kan zijn hond onder controle houden, maar Stijn niet. Honden luisteren nu eenmaal niet naar iedereen.

11. Stijn en Luuk keren steeds weer terug in mijn verhaaltjes. Zij hebben altijd hun mond vol over ‘de arbeider’ en ‘de natuur’. Stijn heeft de overtuiging dat zijn ideeën echt socialistisch zijn, maar Luuk niet. Het is tegenwoordig de vraag of het echte socialisme nog wel bestaat.

12. Luuk en Stijn volgen de cursus ‘Koken voor dommerdjes’. Vandaag moesten zij een appeltaart bakken en hun projectjes kan men het best omschrijven als ‘interessant’. Luuk durft zijn creatie zelfs niet te proeven, maar Stijn wel. Tot nu toe zijn er nog geen doden gevallen tijdens de cursus, maar je weet maar nooit.

13. Anouk en Roos gaan elk jaar snowboarden in Aspen en zijn er dol op. Zij maakten op de laatste skivakantie een sprong over een minuscuul schansje. Anouk vond haar sprong stoer, maar Roos niet. Echt hoog springen, dat kunnen de locals die het hele jaar niets anders doen.

14. Stijn en Luuk kennen elkaars grootouders vrij goed. Zij hebben allebei een beetje een gekke opa en daar denken ze zeer verschillend over. Stijn kan zijn vreemde opa niet waarderen, maar Luuk wel. Waarschijnlijk is het gewoon een vorm van schaamte.

15. Roos en Anouk hebben sinds kort het chatten op het internet ontdekt. Zij slapen sindsdien nog maar maximaal twee uur per dag, maar maken zich daar niet even veel zorgen over. Roos denkt dat haar chat-gedrag niet normaal is, maar Anouk gek genoeg wel. Het valt te hopen dat ze het allebei een beetje binnen de perken kunnen houden.

16. Anouk en Roos zijn allebei ontzettend dol op koffie. Zij maakten afgelopen zondag allebei voor het eerst een koffie verkeerd, met wisselend succes. Van het tweetal vond Anouk haar probeerseltje lekker, maar Roos niet. Misschien kun je in sommige gevallen echt spreken van een koffie verkeerd.

17. Roos en Anouk hebben in de loop der tijd veel gouden sieraden verzameld. Grappig genoeg hebben ze allebei op een bepaald moment in hun leven twee gouden tanden gehad. Roos vond haar gouden tanden ook daadwerkelijk mooi, maar Anouk niet. Gelukkig wordt iedereen op een gegeven moment volwassen.

18. Anouk en Roos spelen samen in een punkbandje. Anouk bespeelt een normale gitaar en Roos de basgitaar. Zij gingen laatst helemaal los tijdens hun solo. Van beide 'gitaarvirtuozen' was Anouk niet tevreden over haar timing en algemene performance tijdens de solo, maar Roos wel. De ene dag gaat het nu eenmaal voor je gevoel wat beter dan de andere.

19. Luuk en Stijn zitten samen op de filmacademie. Zij hebben vandaag allebei de laatste hand gelegd aan hun nieuwste korte film. Van de twee jonge cineasten vindt Luuk dat zijn film goed gelukt is, maar Stijn niet. Uiteindelijk gaat het er natuurlijk wel om wat het publiek ervan vindt.

20. Stijn en Luuk kwamen toevallig tegelijk terug van een lange reis. Zij gaven allebei bij terugkomst een feestje voor hun vrienden, maar waren niet even tevreden met de avond. Van hen beiden vond Stijn zijn feestje niet geslaagd, maar Luuk wel. Soms verheug je je gewoon teveel op iets.

21. Toen ze samen aan het reizen waren door Australië, hebben Roos en Anouk leren surfen. Zij wankelden voortdurend op het surfboard. Roos dacht dat haar surfstijl professioneel aandeed, maar Anouk niet. 'Hang loose' en 'no worries' zijn veelgehoorde kreten op de Australische stranden, maar niet altijd van toepassing op onhandige West-Europese meisjes.

22. Luuk en Stijn zijn net jarig geweest. Zij hebben allebei een nieuwe fiets gekregen, maar zijn er niet even blij mee. Luuk vindt zijn nieuwe fiets mooi, maar Stijn helaas niet. Toch hebben ze afgesproken morgen een leuk tochtje te maken.

23. Anouk en Roos waren allebei op zoek naar een nieuwe computer. Zij hebben zich laten verleiden door een aanbieding bij de Mediamarkt en hebben een laptop aangeschaft. Anouk vindt haar laptop handig, maar Roos absoluut niet. Soms is het beter een wat duurder product te kopen.

24. Stijn en Luuk hebben allebei een grand-slam toernooi gewonnen in 2014. Zij presteren momenteel volgens iedereen in het circuit op het toppen van hun kunnen, maar één van hen heeft daar duidelijk een andere mening over. Stijn denkt toch dat zijn spel nog verbeterd kan worden, maar Luuk niet. Zonder zelfvertrouwen blijf je immers niet nummer één.

Fillers

1. Voor een bruisend studentenleven moet je in Groningen zijn. Veel studenten, waaronder Luuk, besteden dan ook veel te veel tijd aan feesten en lezen niets. Iedere student die net zoals Luuk eigenlijk geen boek aanraakt, weet vrijwel zeker dat hij langer dan vier jaar over zijn studie zal doen. Dat wordt dus lenen!

2. Er woedde een hevige brand in een oude loods. Toen de loods plotseling instortte, raakte brandweerman Stijn helaas gewond. Elke brandweerman die zag hoe Stijn snel werd afgevoerd naar het ziekenhuis, vond het erg jammer dat hij nu niet meer mee kon helpen. Je wilt als brandweerman natuurlijk niet verslagen worden door een brand.

3. Het was een schitterende zomerse dag en druk op het strand. Roos lag lekker te bakken op het warme zand, maar was te lui om zich in te smeren. Roos nam niet de moeite zich te beschermen. Het was al vrij snel te zien dat zij helemaal rood werd. Dat is dus typisch een voorbeeld van: eigen schuld, dikke bult!

4. Wimbledon is het belangrijkste tennistoernooi van het jaar. Voor Anouk is het winnen van dit toernooi dan ook het hoogst haalbare. Anouk, die zich ieder jaar weer perfect voorbereidt, acht het mogelijk dat zij dankzij deze volledige inzet ooit eens zal triomferen. Een tenniscarrière kan

daarna niet meer mislukken.

5. Het grote dorpsfeest is weer in aantocht. Veel boeren uit de omgeving, waaronder Stijn, laten de koe dan even de koe. Elke boer die zich net als Stijn tegoed zal doen aan bier en worst, weet eigenlijk nu al dat hij uiteindelijk ladderzat met de tractor naar huis gaat. Gelukkig gaat een beetje tractor niet veel harder dan dertig kilometer per uur.

6. De belangrijke finale ging weer eens tussen Ajax en PSV. Helaas raakte Luuk, de aanvoerder van Ajax, geen één bal goed. Iedere Ajax-speler die zag hoe Luuk toch voortdurend zijn best deed, wenste natuurlijk wel dat hij snel beloond zou worden voor deze inzet. Anders zou er zeker een kleedkamertje sneuvelen.

7. Een werkweek bestaat eigenlijk voornamelijk uit het uitzien naar zaterdag. Anouk zal dat zeker beamen. Anouk droomt al op maandag over het weekend. Het is natuurlijk vrij duidelijk dat zij dan lekker lang in bed blijft liggen. Maar wat duurt een week toch lang!!

8. Gedurende een schaakwedstrijd worden de stukken soms achteloos verschoven. Vooral Roos lijkt dan heel snel alle mogelijkheden te berekenen. Roos, die computerachtige trekjes lijkt te vertonen, geeft daarentegen de verklaring dat zij gewoon bepaalde stellingen herkent. Als je een stelling herkent, weet je natuurlijk veel sneller wat je moet doen.

9. Onze samenleving is duidelijk aan het veranderen. Eerst vonden bijna alle Nederlanders, zoals Anouk, geld het belangrijkste in hun leven. Iedereen die net als Anouk vooral een dikke vette bankrekening wilde, ziet nu gelukkig in dat zij uiteindelijk kapot zal gaan aan dat streven. Geld is een middel, geen doel op zich.

10. Thailand staat bekend om zijn tempels en zwoele zomeravonden. Roos heeft dan ook Thailand bezocht om cultuur te snuiven en lekker te luieren. Iedereen die weet dat Roos ruim een half jaar in dit land heeft rondgezworven, merkt zonder enige twijfel dat zij gedurende haar verblijf duidelijk een andere kijk op het leven heeft gekregen. De mysterieuze Thaise glimlach zorgt daar wel voor.

11. Zoals elk jaar werd in december 2003 de Gouden Loeki uitgereikt voor het beste reclamespotje. Luuk was genomineerd en aanwezig tijdens de grootse uitreiking. Luuk behoorde

tot de echte kanshebbers. Het was eigenlijk zeker dat hij die avond de bokaal mee naar huis zou nemen. Die zou vast en zeker mooi staan op de schoorsteenmantel.

12. Verliefd zijn heeft dezelfde werking als een sterke drug. Stijn is verliefd en zal dat zeker beamen. Stijn, die momenteel in een droomwereld leeft, weet echter maar al te goed dat hij ooit weer met beide benen op de grond zal staan. Dat is maar goed ook, want anders zou er van werken weinig terechtkomen.

13. Het is tegenwoordig maar moeilijk overleven op een middelbare school. Zo moeten docenten, zoals bijvoorbeeld Roos, erg hun best doen om populair te zijn. Iedere docent die net als Roos probeert cool over te komen, is wel een beetje bang dat zij daardoor minder autoriteit uitstraalt. Maar wat moet je anders, als streng zijn geen enkel effect heeft?

14. Het heeft ook nadelen om ontzettend beroemd te zijn. Zo heeft de muzikante Anouk de grootste moeite haar leven op de rails te houden. Iedere muzikant die weet dat Anouk niet goed tegen de druk kan, is ontzettend bang dat zij continu verslaafd raakt aan het één of ander. De verleidingen van drugs en drank zijn nu eenmaal groot.

15. Economisch gezien gaat het niet zo lekker in Nederland. Zo heeft Stijn het als kleine ondernemer erg moeilijk. Stijn moet elk dubbeltje omdraaien. Het is inmiddels wel duidelijk dat hij waarschijnlijk failliet zal gaan. Op dit moment kan alleen bidden nog helpen.

16. Vreemdgaan is al zo oud als de mensheid. Luuk kan daarover meepraten. Luuk, die vaak in de verleiding komt, hoopt over het algemeen dat hij sterk en dus trouw zal zijn als de volgende kans zich aandient. Monogamie staat in onze samenleving immers nog steeds hoog in het vaandel.

17. De mensheid wordt al sinds haar ontstaan gefascineerd door vliegen. Piloten, zoals Luuk, zijn daar natuurlijk een goed voorbeeld van. Iedere piloot die net als Luuk kort geleden zijn vliegbrevet heeft gehaald, voorspelt met veel zekerheid dat hij nooit de dag van zijn eerste vlucht zal vergeten. Het valt namelijk te vergelijken met een engel die zijn vleugels krijgt.

18. In feite is de wetenschap een nieuwe religie die enige zin geeft aan het leven. Wetenschapper Stijn is in zo'n vergelijking een trouwe volgeling van een godsdienst. Elke wetenschapper die

weet dat Stijn dag en nacht bezig is met dit nieuwste geloof, is ervan overtuigd dat hij daarmee bijdraagt aan de vooruitgang van de mens. Het blijft natuurlijk wel de vraag wat vooruitgang nu precies is.

19. De derde etappe van de Tour de France stond vandaag op het programma. Roos was toeschouwer en zag aan de donkere lucht dat het snel ging regenen. Roos had daar niet op gerekend. Het was natuurlijk dom dat zij geen paraplu had meegenomen. Als je een hele dag van huis bent, moet je op regen voorbereid zijn.

20. De politie is vaak sneller op de plaats van een medisch spoedgeval dan de ambulance. Daarom leert Anouk als politieagente momenteel belangrijke eerste-hulp technieken. Anouk, die hard moet leren voor het ingewikkelde examen, vertelt met veel overtuiging dat zij graag tijd vrijmaakt om deze vaardigheden onder de knie te krijgen. Levens redden is immers een nobele bezigheid.

21. Het is een bekend gegeven dat bij het maken van een film nooit al het materiaal gebruikt wordt. Toch waren veel actrices, waaronder Roos, erg teleurgesteld tijdens de première van hun film 'Anarchie'. Iedere actrice die net als Roos vol ongeloof naar het uiteindelijke resultaat keek, had namelijk door dat zij minstens uit tien scènes was geknipt. Dat was een beetje teveel van het goede.

22. Morgen begint het jaarlijkse padvindersfestijn. Luuk zal dan als nieuwe padvinder hutten bouwen en oude vrouwtjes helpen oversteken. Iedere padvinder die weet dat Luuk van de spanning haast niet kan slapen, verwacht zonder enige twijfel dat hij een ontzettend leuke tijd tegemoet gaat en na afloop een lintje krijgt. Een padvinder zonder lintje is namelijk zoiets als een piraat zonder ooglapje.

23. Kansspelletjes lijken eigenlijk heel onschuldig. Waarom kan gokverslaafde Stijn dan niet stoppen met gokken? Stijn verbrast soms tienduizenden euro’s. De verklaring is dat er een fysieke pijn ontstaat, als hij niet regelmatig gokt. Er is dus echt sprake van een lichamelijke afhankelijkheid.

24. Het is soms niet echt prettig om in de keuken van een goed restaurant te werken. Anouk, de chefkok van het Hilton, duldt namelijk geen enkele tegenspraak. Anouk, die autoriteit hoog in het vaandel heeft staan, stelt zelfs overduidelijk dat zij de keuken beschouwt als haar koninkrijk.

Daar moet geluisterd worden naar 'de koningin'.

Practice Items

1. Anouk en Roos zitten in hun laatste studiejaar. Zij hebben net met veel pijn en moeite hun scriptie afgerond, maar ze zijn niet even gelukkig met het resultaat. Anouk is tevreden over haar scriptie, maar Roos spijtig genoeg niet. Het is blijkbaar toch wel erg moeilijk om een goed verhaal op papier te krijgen.

2. Roos en Anouk coachen de twee hockeyteams die momenteel op het veld staan. Plotseling ontstaat er een grote vechtpartij tussen de twee teams, maar ze grijpen niet allebei direct in. Roos is helaas de enige die haar speelsters tot de orde roept. Het wordt echt een drama!

3. Anouk en Roos bereiden zich voor op de Olympische Spelen van Athene. Anouk zit in een behoorlijk dipje, of zeg maar gerust een dip! Roos is de enige van de twee topsportsters die verwacht dat haar topvorm nog op tijd gaat komen om een kans te maken op een medaille. Je moet nu eenmaal in supervorm zijn wil je op het podium belanden.

4. Luuk en zeker zijn manager Stijn zijn ontzettend blij dat Luuk weer mag gaan racen. Voor het nieuwe team rijdt Luuk in een felle oranje F1-auto. Luuk vindt zijn nieuwe wagen mooi, maar Stijn helaas niet. Maar eigenlijk is het het belangrijkste dat er eindelijk weer wat geld verdiend wordt om al die verkeersbonnen een keer te betalen.

5. Stijn en Luuk waren blij toen de Tweede Wereld Oorlog was afgelopen. Toch moest Stijn waarschijnlijk een beetje uitkijken, want hij was het neefje van Hitler. Stijn dacht dat zijn achternaam niet voor problemen zou zorgen, maar Luuk wel. Wat heb je toch een naïeve mensen op deze wereld.

6. De rappers Luuk en Stijn brengen tegelijkertijd hun nieuwst elpee uit. Zij lijken allebei vol vertrouwen, zoals altijd. Maar van de twee popsterren denkt Luuk dat zijn nieuwe album het goed gaat doen en Stijn eigenlijk niet. Eerlijk gezegd valt het gewoon niet te voorspellen.

7. Roos en Anouk treden samen op in een nieuwe show. Roos gedraagt zich soms alsof zij de wereld bezit en is daar zelfs trots op. Als enige schaamt Roos zich nooit voor haar diva-gedrag. Roem kan blijkbaar vreemde dingen met je doen.

8. De favoriete bezigheid van Anouk en Roos is toch wel knippen en plakken. Vooral Anouk maakt vaak een creatief doch plakkerig kunststuk. Anouk denkt dat haar werk ooit in het Stedelijk museum zal hangen, maar Roos niet. Het zal allemaal sowieso nog wel een tijdje duren, aangezien ze pas twintig jaar oud zijn.

9. Het is zover, Luuk en Stijn gaan nu echt samen fietsen in Tibet. De plannen die ze maken, worden per week wilder en wilder. Luuk denkt als enige dat zijn conditie goed genoeg is om het basiskamp op de Mount Everest te bereiken. Ze zullen daar raar opkijken als iemand op een fietsje aan komt rijden.

10. Zwemmen op de Zuidpool is het coolste wat Stijn en Luuk in hun leven hebben gedaan. Een beetje hachelijk was toch wel het moment dat Stijn met een orka aan het spelen was. Stijn had niet door dat zijn leven in gevaar was, maar Luuk wel. In het Engels heten orka's heten niet voor niets "killer whales".