INFLUENCE OF VISUAL AND LINGUISTIC CONTEXT ON OBJECT PRONOUN PROCESSING: EEG AND EYE-TRACKING PROVIDE NEW INFORMATION

Bachelor's Project Thesis

Rosa Verhoeven, s2678152, r.verhoeven.1@student.rug.nl

Supervisors: Dr J.P. Borst & Dr J.C. van Rij-Tange

Abstract: This study aimed to investigate whether and how visual and linguistic context affect the processing of object pronouns ('him', 'her'). Recent studies have shown, by measuring pupil dilation, that context information affects object-pronoun processing in an early stage and interacts with grammatical processing. In the current study, we investigated the influence and the exact timing of those contexts on object-pronoun processing by co-registering EEG and pupil dilation. With a 2x2x2 within-subject design, we investigated the effects of visual context (other-oriented vs self-oriented action; e.g., a picture with a mouse touching a squirrel with a spoon, or a mouse touching himself with a spoon), discourse prominence (the introduction sentence introduces the actor first or second; '*You just saw a mouse and a squirrel*.' vs '*You just saw a squirrel and a mouse*.'), and referring expression (the test sentence contains an object pronoun or a reflexive (fillers); 'The mouse touched him / himself with a spoon.'). Surprisingly, we found that visual and linguistic context do not have an influence on object-pronoun processing. However, we did find an influence of visual context on the processing of reflexives in EEG and an interaction between visual and linguistic context in pupil dilation.

1. Introduction

In languages such as Dutch and English, the interpretation of object pronouns (*"him", "her"*) is guided by Principle B of Binding Theory (Chomsky, 1981). That is, in sentence pairs like *"You just saw a mouse and a squirrel. The mouse touched him with a spoon"*, the object pronoun *him* cannot refer to the same individual as the subject of the local clause, being the hedgehog in this case. Therefore, it is interpreted by adults as referring to another referent in the context, which would be the mouse.

Two conflicting types of accounts have been proposed to explain the time course of object pronoun resolution. According to the *initial-filter account*, the previously mentioned grammatical restriction plays the foremost role in the interpretation of object pronouns. It assumes that other sources of information such as linguistic discourse only play a minor role, if any (e.g., Nicol & Swinney, 1989; Clifton, Kennison, & Albrecht, 1997; Chow, Lewis, & Phillips, 2014). However, several recent studies have suggested that context information does affect object pronoun processing in an early stage, thereby supporting the *competing-constraints account* which assumes that other sources of information compete with grammatical constraints (a.o., Badecker & Straub, 2002; Kennison, 2003).

One of these studies (van Rij, 2012), reports a combined effect of visual context (i.e., visual scenes being congruent with the pronoun interpretation or not) and linguistic context (i.e., discourse prominence, the order of introduction of the referents) on pupil dilation 500-1000 ms after the onset of the object pronoun. The results suggest that with a canonical introduction order (i.e., the actor being introduced first), a larger pupil dilation is elicited in case the participant's interpretation of the pronoun and the scene are incongruent, in comparison to the congruent situation. Moreover, a canonical introduction order with a congruent scene elicits a smaller pupil dilation than both the congruent and incongruent scenes with a non-canonical introduction order (in which the actor is mentioned second).

The reason for the influence of the order of mention is that first mentioned referents are perceived as more prominent, hence more likely antecedents of a subject pronoun than referents that are mentioned later (e.g., Gernsbacher &

Hargreaves, 1988; Gordon, Grosz, & Gilliom, 1993; Kaiser & Trueswell, 2008). That is, people expect a first mentioned referent to be mentioned again as the subject of the following sentence, and the other referent to be mentioned again in a less prominent grammatical position (for example, using an object pronoun). If the subject of the second sentence does indeed refer to the first mentioned referent, the listener will take the picture as the correct interpretation of the sentence. However, if the subject of the second sentence does not refer to the referent that had been mentioned first in the previous sentence, but to the one that had been mentioned later, there is conflicting information about who is important. This may cause listeners to be careful and wait for the object before building a sentence representation. Hence, when the referents are introduced in a non-canonical introduction order, the listener is less surprised when the scene turns out to be incongruent and directly knows whether the scene was congruent or not, whereas with a canonical introduction order, the listener has already built up an expectation based on the visual scene and is thus more surprised when the scene was incongruent, leading them to have to revise their interpretation.

Pupil dilation is sensitive to subtle linguistic differences, associated with differences in processing load (e.g., Engelhardt, Ferreira, & Patsenko, 2010; Zellin, Pannekamp, Toep, & van der Meer, 2011). Whereas it is a quite effective measure, it is relatively slow, given that, on average, the dilation peaks only 1000 ms after the trigggering stimulus (e.g., Hoeks & Levelt, 1993). Hence, it does not provide precise information on the timing of the effects of visual and linguistic context. In that regard, recording the elektroencephalogy (EEG) signal may provide more temporally informative results.

The aim of this study was therefore to investigate the influence and the exact timing of the effect of visual and linguistic context on object pronoun processing by co-registering EEG and pupil dilation.

In contrast with earlier studies, the current study's design consisted of a picture verification task in combination with a blank screen paradigm (Altmann, 2004). First, a visual scene was presented followed by a blank screen during which two prerecorded, Dutch sentences were played like the earlier mentioned mouse-squirrel example. Native Dutch participants then had to judge whether these sentences were a correct description of the scene or not. The benefit of this paradigm is that it reduces the number of eye-movements during sentence processing, resulting in fewer artifacts. Moreover, it allowed us to distinguish activity caused by sentence processing from activity caused by looking at the scene.

1.1. Research question

The current study aimed to answer the following research question:

Do visual and linguistic context (introduction order) have an influence on object pronoun processing in a blank screen paradigm? If there are effects:

- What are the effects on pupil dilation?
- Can these effects be seen in the EEG?
- When do they occur?

1.2. Hypothesis

This study has tested the hypothesis that both visual context and linguistic context set up an expectation for an interpretation of the object pronoun. When this interpretation was not congruent with the actual pronoun that had been presented in the prerecorded sentences, we expected to see a higher N400 amplitude compared to when it was congruent. We expected to see this effect of congruency especially when the order of introduction was canonical (more difficult retrieval, according to Brouwer & Hoeks, 2013; or violating expectations, according to van Berkum, Korneef, Otten, & Nieuwland, 2007). We expected to see less of a difference between the introduction orders for the processing of pronouns with congruent pictures.

However, when the incongruence is only detected after grammatical processing like the initial-filter account assumes, we would not expect to see such an early effect; instead, we may see later amplitude differences, such as in P600 (associated with syntactic and pragmatic integrative processes, according to Brouwer & Hoeks, 2013).

In addition to these results, we expected the pupil dilation to be in agreement with previous studies, allowing for a comparison with earlier pupil dilation studies (van Rij, 2012) and the new EEG results.

2. Method

2.1. Participants

In total, 32 students (right-handed, native Dutch speakers) participated in our experiment of whom 18 were male and 14 female. The average age of the participants was 21.5 years old, ranging between the ages of 19-24. All participants signed an informed consent form before the start of the experiment. On average, the total time spent on both setting up and conducting the experiment took approximately 1.5 hours. Participants received a compensation of €12,-.

2.2. Design

By combining a picture verification task with a blank screen paradigm, we investigated the influences of both visual and linguistic context on object and reflexive pronoun processing.

The 2x2x2 (Picture Type x Introduction Order x Referring Expression) design was tested within subjects and partly within items (four variants of each item were tested in each experimental session).

For Picture Type there were two conditions: other-oriented pictures (see Figure 2.2.1), in which the actor (an animal) is performing an action upon another animal, and self-oriented pictures (see Figure 2.2.2), in which the actor is performing the action upon himself. There were 80 unique pictures, which formed 40 pairs of two variants of the same picture.

Introduction Order is related to the introduction sentence which followed the picture. There were two conditions: an actor-first introduction, in which the actor is mentioned first (“*Zojuist zag je een muis en een eekhoorn.*”, you just saw a mouse and a squirrel) or actor-second (“*Zojuist zag je een eekhoorn en een muis.*”, you just saw a squirrel and a mouse).

After the introduction sentence, the test sentence was played, containing the referring expression. This could be either an object pronoun, “*hem*”, (“*De muis raakte hem aan met een lepel*”, the mouse touched him with a spoon) or a

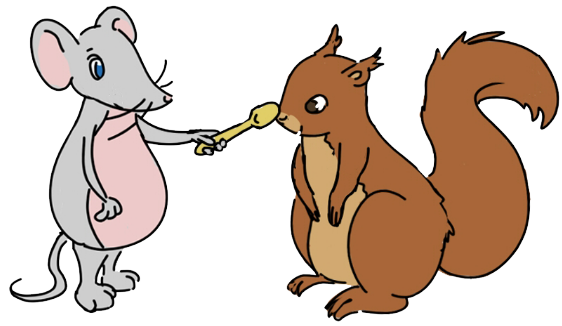

Figure 2.2.1: an example of an other-oriented visual stimuli

Figure 2.2.2: an example of a self-oriented visual stimuli

reflexive pronoun, “*zichzelf*” (“*De muis raakte zichzelf aan met een lepel*”, the mouse touched himself with a spoon). The reflexive sentences were included as fillers.

Each picture was shown twice to the participants, with each image occurring in a different block and with two of the four conditions (Introduction Order x Referring Expression). We tested multiple variations of each item to increase the amount of data necessary for EEG analysis. To avoid any repetitions of these combinations, or to avoid two of the same combinations after each other, we made unique lists for all participants. These lists consisted of four different blocks. For each participant, the order of items within each block was randomized to avoid any further bias.

2.3. Material/Stimuli

Stimulus presentation was programmed in Experiment Builder (SR Research, 2017).

The pictures consisted of a combination of the visual stimuli of van Rij (2012) and van Rij et al (2016). They were presented centrally against a light grey background with a width of 500 pixels.

The height depended on the image ratio. 50% of the pictures were randomly selected and mirrored.

For each picture, we recorded a two-sentence description. These sentences were recorded in the recording studio of the Faculty of Arts, University of Groningen, and were afterwards manipulated, by means of splicing and normalizing, with the program PRAAT (Boersma & Weenink, 2018).

Two kinds of introduction sentences were recorded: actor-first and actor-second. They were all built in a similar style with artificial breaks: "*Zojuist zag je*" (you just saw) + 100 ms silence + <referent> + 100 ms silence + "*en*" (and) + <referent>.

Three kinds of test sentences were recorded: with either an object noun sentence, an object pronoun or a reflexive pronoun. The sentence with an object noun sentence (e.g., *"De muis raakte* de eekhoorn *aan met een lepel"*, the mouse touched *the squirrel* with a spoon) was used as carrier phrase for the test sentences. The pronoun ("*hem*") and reflexive ("*zichzelf*") were spliced into these object noun parts, so the intonation of the rest of the sentence would be kept identical. The test sentence, too, had artificial breaks: <Actor> + 100 ms silence + <verb> + 100 ms silence + <pronoun/reflexive> + 100 ms silence + <prepositional phrase>.

Between the introduction sentence and the test sentence was a fixed break of 200 ms.

The answer screen contained two boxes: one green, with the word "correct" in it, and one red, with the word "incorrect" in it. The order of these boxes was randomly determined for all trials in the experiment to prevent motor preparation. The Ctrl-left button was linked to the left-positioned answer, the Ctrl-right button to the right.

The pupil of the left eye was monitored continuously during the picture verification task with the EyeLink 1000 (SR research) at 500 Hz (16 mm lens + target sticker). Brain activity was measured via EEG caps consisting of 32 electrodes and six external electrodes: on the mastoids, HEOG and VEOG (above and below the right eye). These were connected to BioSemi, which recorded the data at 2048 Hz.

2.4. Procedure

In advance, the participant was informed to wear neither glasses nor mascara, since both influence the precision of the eye-tracker's pupil detection - hence, the eye-tracking results. Before going on to the actual experiment, the participant had to sign the consent form. The participant was positioned on a non-adjustable chair behind a computer screen, which was positioned on a desk. On the ground, the target position for the chair was indicated using tape, to make sure all participants would be facing the eye-tracking from a fixed location. The eye-tracker was installed in front of the computer, at a distance of approximately 70 centimeters to the participant's eyes. The keyboard was positioned between the eye-tracker and the participant, at a distance comfortable for the participant.

After this part of the set-up, the EEG cap and electrodes were positioned. On the participant's forehead, we placed a target sticker (i.e., a sticker with a bullet point on it), for the eye-tracker to detect as target point.

An oral instruction was followed by a similar instruction on the screen: during this instruction, participants became familiar with the kind of pictures they were about to face and the keys they had to press accordingly. Then, they had to perform an eye-tracker calibration followed by a validation. We aimed for an average deviation value of 0.5. If it was higher than 0.5, another calibration had to be conducted (the lower this value, the more precise the calibration and thus the eye-tracking data).

Figure 2.4.1 visualizes the structure of the trial. Each trial started with a fixation point. An invisible square surrounded this fixation point and the participant had to look for at least 100 ms within this square to start with the trial. 650 ms after this, the picture appeared on the screen, and was shown for 2000 ms. A blank screen followed. 500 ms later, while still seeing the blank screen, the two-sentence stories were played. 1200 ms after the offset of the test sentence, the answer screen appeared during which the participant had to indicate whether the story was congruent with the picture or not. They had 5 seconds to give an answer.

If the eye tracker did not recognize that the

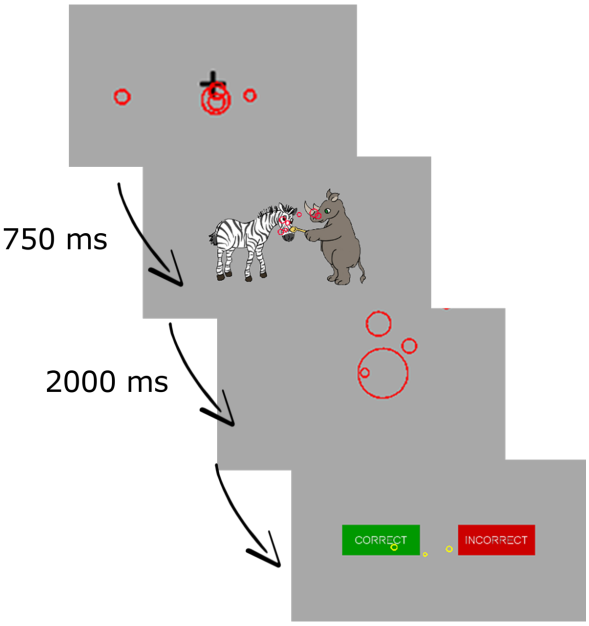

Figure 2.4.1: Example of a trial with a participant's fixation points

participant was looking within the invisible square surrounding the fixation point for 100 ms within 5 seconds, another calibration had to be performed and the trial would be skipped.

The participant started with three practice trials with pictures that were not used during the actual experiment.

After the participant had finished the practice trials and asked all their possible questions, they would perform another calibration before continuing to the actual experiment. In total, the participant performed 160 trials, divided into 4 blocks of 40 trials. Each block was separated from the next one by a break. After each break, the participant was asked to calibrate again.

Luminance of the room was normal and kept constant during the experiment.

2.5. Analysis methods

The EEG data has been pre-processed with a script from Jelmer Borst, using EEGLAB (Delorme & Makeig, 2004). The data has been re-referenced to the mastoids electrodes and downsampled to 100 Hz. The low pass filter was set to 40 Hz, which removes fast noise. The high pass filter was set to 0.01 Hz, to remove very slow noise. After the downsampling and filtering, trials with extreme values were manually rejected. Blinks and saccades have been removed with ICA.

The pupil dilation data has been automatically pre-processed with a script from Jacolien van Rij, using R (R Core Team, 2018). Blinks and saccades have been removed from the data, with 100 ms padding around the blink and 10 ms around the saccade.

The first 80 trials (that is, the first two blocks) have been used for the analysis, since then the participants had only encountered each picture once and these results were more reliable as the participants reported they became more distracted during the last two blocks.

The data was baselined on 250-0 ms before the pronoun onset in order to investigate differences between conditions starting from the pronoun onset.

Statistical analyses have been performed on the eye-tracking and EEG data, using linear mixed-effects (LME) models. Even though the reflexive sentences were only meant as fillers, we have performed analyses on both sentences types.

For eye tracking, the window on which the analyses have been performed is 750 to 1250 ms after the onset of the pronoun. The reason for this is that pupil dilation peaks around 1000 ms after the stimulus onset that triggered the dilation. For each subject, the median pupil size per trial within this window has been taken and analyzed.

In EEG, we expected to see a N400 or a P600 when the trial is incongruent. Therefore, we performed an analysis on two time windows: the first from 300 to 500 ms after the pronoun onset and another from 500 to 700 ms after the pronoun onset. Similar to the pupil dilation, the median Cz value has been taken for all the subjects per trial, followed by a comparison of the means of those.

For all of the data, three models have been assessed: the simplest one only checking for main effects, the second also including all two-way interactions and the most complex one adds the three-way interaction to that. These two latter, more complex, models have only been used if they proved to explain significantly more variation than the simplest model. If the most complex model, which includes the three-way interaction, turned out to be the best model, two separate analyses have been performed on the two sentence types (pronoun and reflexive), apart from each other.

3. Results

This section will start with an analysis of the influences of the different conditions in accuracy and reaction times. This will be followed by an analysis on the linear mixed effect models of pupil dilation and then a similar analysis on the EEG data. Finally, a comparison will be made between the data of the pupil dilation and EEG.

3.1. Accuracy and reaction time

During the experiment, the accuracy and reaction times for all of the four conditions (Picture Type x Introduction Order) for both the pronoun and the reflexive were measured. Figure 3.1.1 shows the accuracy and Figure 3.1.2 shows the response times for each condition. In both figures, 'Other' and 'Self' represent the two options for Picture Type: other-oriented and self-oriented, respectively. A1 and A2 represent the two options for Introduction Order: actor-first and actor-second, respectively.

On average, the accuracy is high (96.8%). Moreover, there is no difference in accuracy between the different conditions. This shows that the participants understood our experiment very well and that the conditions did not have an influence on how well the participants performed the experiment.

The average response time is 604.6 ms. The different conditions do not have an influence on the reaction time, either.

3.2. Pupil dilation

Figure 3.2.1 shows the results of the pupil dilation for the pronoun trials, in which 0 ms represents the onset of the object pronoun (i.e., the beginning of the word '*hem*'). The x-axis represents time in ms and the y-axis represents the pupil dilation in arbitrary units, chosen by the EyeLink 1000. The kind of Picture Type is represented by color: the black lines represent other-oriented (O) pictures, whereas the red lines represent self-oriented (S) pictures. Introduction Order is represented by the style of the line: the solid lines represent an actor-first (A1) introduction and the dashed lines represent an actor-second (A2) introduction.

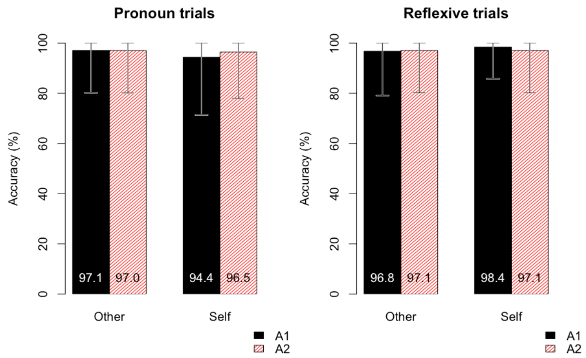

Figure 3.1.1: Accuracy for each condition

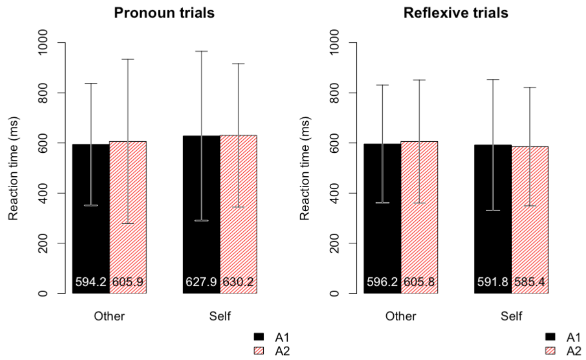

Figure 3.1.2: Response times for each condition

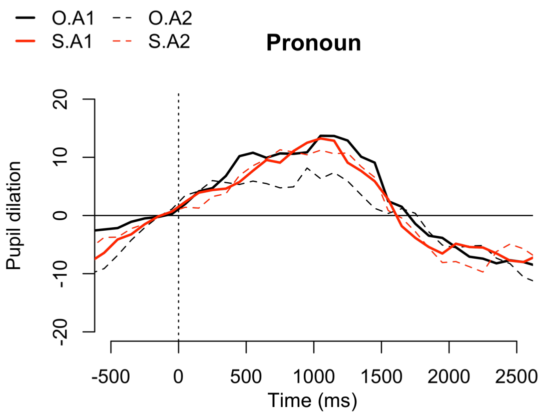

Figure 3.2.1: Pupil dilation in object pronoun

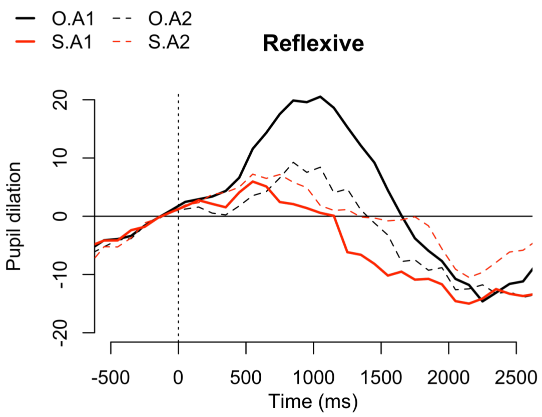

Figure 3.2.2: Pupil dilation in reflexive pronoun

From the looks of the graph, there does not seem to be a difference between the different conditions (that is, Picture Type x Introduction Order) in object pronoun processing.

The results of the pupil dilation for the reflexive pronoun can be seen in Figure 3.2.2. The other-oriented (thus incongruent), actor-first condition seems to elicit a larger pupil dilation than the other conditions. This indicates an interaction effect between Picture Type and Introduction Order on reflexive pronoun processing.

A backward-fitting model comparison procedure has been performed on three linear mixed effect models: the simplest model only considering main effects, then a model which also includes two-way interactions and the most complex model adding the three-way interaction to that. The most complex model explains significantly more variance than the model which only contains the two-way interactions ($X^2(1) = 5.740$; $p = 0.017$). The summary of this model is shown in Table 3.2.1. It shows a significant main effect for Sentence Type ($\beta = 17.684$, SE = 7.146, t = 2.475). Moreover, it shows that the two-way interaction Picture Type x Sentence Type is significant ($\beta = -41.064$, SE = 10.057, t = -4.083). Finally, it shows a significant effect of the three-way interaction Picture Type x Introduction Order x Sentence Type. ($\beta = 34.175$, SE = 14.255, t = 2.397). This suggests that the interaction between Picture Type and Introduction Order is influenced by Sentence Type. Therefore, the data will now be split into two categories: one for the object pronoun, one for the reflexive pronoun.

Table 3.2.1: Fixed effects on pupil dilation of the model with three-way interaction

Model: medianPupil ~ (introtype + pictype + sentencetype)^3 + (1 | Subject) + (1 | Item)			
	Estimate	**Std. Error**	**t value**
(Intercept)	17.905	6.175	2.900
introtypeA2	-5.064	7.127	-0.711
pictypeS	8.478	7.140	1.187
sentencetypeR	17.684	7.146	2.475
introtypeA2: pictypeS	-1.458	10.095	-0.144
introtypeA2: sentencetypeR	-18.022	10.116	-1.782
pictypeS: sentencetypeR	-41.064	10.057	-4.083
introtypeA2: pictypeS: sentencetypeR	34.175	14.255	2.397

Table 3.2.1.1: Fixed effects on pupil dilation of the model with main effects for pronouns

Model: medianPupil ~ (introtype + pictype) + (1 | Subject) + (1 | Item)			
	Estimate	**Std. Error**	**t value**
(Intercept)	18.333	5.809	3.156
introtypeA2	-5.854	5.251	-1.115
pictypeS	7.725	5.250	1.471

Two models have been made for each of those categories: one with only the main effects, and one which also contains the two-way interaction Picture Type x Introduction Order. Backward-fitting model comparison procedures have been performed on these models.

3.2.1. Pupil dilation: pronoun analysis

For the pronoun, the more complex model with two-way interaction does not explain significantly more variance than the simpler model which only tests for main effects ($X^2(1) = 0.018$; $p = 0.894$). Therefore, the latter is used for the analysis.

Table 3.2.1.1 shows the summary of the model. It shows that the main effects, Picture Type and Introduction Order, do not have an effect on pronoun processing. This is in line with the claims about Figure 3.2.1, which was that the pupil dilation seems to be the same for all of the conditions.

3.2.2. Pupil dilation: reflexive analysis

The same model comparison procedure as to the object pronoun analysis has been performed to the sentences with a reflexive pronoun. The more complex model, which contains the two-way interaction Picture Type x Introduction order, explains significantly more variance than the model without this interaction ($X^2(1) = 11.097$; $p = 0.001$).

The summary of this model, shown in Table 3.2.2.1, explains the influence of multiple effects.

First, there is a significant main effect of Introduction Order on the processing of reflexive pronouns ($\beta = -23.140$, SE = 6.957, t = -3.326 for an actor-second introduction). This means that the actor-second introductions elicit a smaller pupil dilation than the actor-first sentences for other-oriented pictures in reflexive pronoun processing. This can also be observed in Figure

Table 3.2.2.1: Fixed effects on pupil dilation of the model with two-way interaction for reflexives

Model: medianPupil ~ (introtype + pictype)^2 + (1 ∣ Subject) + (1 ∣ Item)			
	Estimate	**Std. Error**	**t value**
(Intercept)	35.663	5.536	6.442
introtypeA2	-23.140	6.957	-3.326
pictypeS	-32.672	6.865	-4.760
introtypeA2: pictypeS	32.603	9.761	3.340

3.2.2: the other-oriented, actor-second condition is lower than the other-oriented, actor-first condition.

Moreover, there is a significant main effect of Picture Type (β = -32.672, SE = 6.865, t = -4.760 for self-oriented pictures), which indicates that the other-oriented (thus incongruent) pictures elicit a bigger pupil dilation than self-oriented (congruent) pictures with an actor-first introduction for reflexive pronoun processing. This, too, is shown in Figure 3.2.2: the other-oriented, actor-first condition is higher than the self-oriented, actor-first condition.

Finally, the interaction between Picture Type and Introduction Order is significantly important (β = 32.603, SE = 9.761, t = 3.340 for an actor-second introduction with a self-oriented picture). This suggests that there is an interaction effect, resulting in a stronger effect of introduction order for other-oriented pictures than for self-oriented pictures. Figure 3.2.2 shows this as well: the other-oriented pictures are much further apart than the self-oriented pictures.

3.3. EEG

Figure 3.3.1 shows the results of the electrical brain activity for the object pronoun, measured with the EEG's Cz electrode. The x-axis again represents time in ms but has a shorter range than pupil dilation as its peaks earlier after stimulus onset. The y-axis represents the amplitude in electrical activity in microvolts.

Similar to the results of the pronoun in pupil dilation, neither Picture Type nor Introduction Order nor the interaction between those two seems to have an influence on pronoun processing in EEG. Both around the N400 and around the P600, there does not seem to be a difference between the conditions.

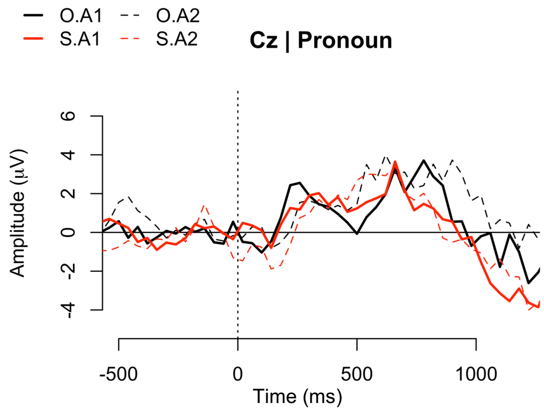

Figure 3.3.1: EEG results for object pronoun

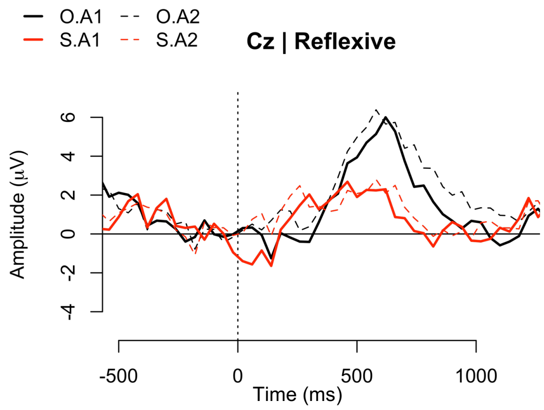

Figure 3.3.2: EEG results for reflexive pronoun

Figure 3.3.2 shows the EEG results for the reflexive pronoun. Around the N400, the different conditions do not seem to be elicit a difference in amplitude. However, in contrast to the results of pupil dilation for the reflexive pronoun, the other-oriented, thus incongruent, pictures, seem to elicit a P600. This indicates an effect of Picture Type on reflexive pronoun processing.

3.3.1. Analysis on the N400

Since we expected to see an influence of visual context and linguistic context on the N400, a backward-fitting model comparison procedure has been performed on linear mixed effect models of this area, first. The most complex model with the three-way interaction does not explain significantly more deviance than the model containing only two-way interactions ($X^2(1) = 0.005$; $p = 0.9445$). This latter model also does not explain significantly more variance than the main effects model ($X^2(3) = 0.8718$; $p = 0.8322$). Therefore, the main effects model is used.

Table 3.3.1.1: Fixed effects on N400 of the model with main effects

Model: Cz ~ (introtype + pictype + sentencetype) + (1 ∣ Subject) + (1 ∣ Item)			
	Estimate	**Std. Error**	**t value**
(Intercept)	1.57271	0.80901	1.944
introtypeA2	-0.08103	0.65172	-0.124
pictypeS	-0.04663	0.65188	-0.072
sentencetypeR	0.56103	0.65518	0.856

Table 3.3.1.1 shows the summary of this model. It shows that neither of the conditions has an influence on the processing of pronouns. This suggests that both for object pronouns and reflexive pronouns, there is no significant influence of Picture Type or Introduction Order around the location of the N400. This is in agreement with what has been observed in Figures 3.3.1 and 3.3.2: from 300-500 ms after the onset of the pronoun, there is no difference in amplitude between the different conditions.

3.3.2. Analysis on the P600

A backward-fitting model comparison procedure has been performed on the area around the P600 as well. The three-way interaction model does not explain significantly more variance than the two-way interaction model ($X^2(1) = 0.123$; $p = 0.726$). Moreover, the two-way interaction model does not explain more variance than the main effects model ($X^2(3) = 5.235$; $p = 0.155$).

However, when looking at the summary of the two-way interaction model (which is slightly, but not significantly, better than the main effects model) in Table 3.3.2.1, the interaction between Picture Type and Sentence Type appears to be significant ($\beta = -3.311$, SE = 1.475, t = -2.243 for a self-oriented picture and a reflexive pronoun). This suggests that the self-oriented pictures with the object pronoun elicit a higher P600 than in the reflexive pronoun. Therefore, we added this single interaction to the model of the main effects. A model comparison has been performed between this model, with main effects and the interaction Picture Type x Sentence Type, and the simpler model with only the main effects. The more complex model with the two-way interaction explains the variance significantly better than the main effects model ($X^2(1) = 4.974$; $p = 0.026$). Therefore, this model has been used.

Table 3.3.2.1: Fixed effects on P600 of the model with all two-way interactions

Model: Cz ~ (introtype + pictype + sentencetype)^2 + (1 ∣ Subject) + (1 ∣ Item)			
	Estimate	**Std. Error**	**t value**
(Intercept)	2.04204	1.12164	1.821
introtypeA2	1.37267	1.27535	1.076
pictypeS	0.08508	1.28682	0.066
sentencetypeR	3.04516	1.27900	2.381
introtypeA2: pictypeS	-0.21279	1.46014	-0.146
introtypeA2: sentencetypeR	-0.71464	1.46041	-0.489
pictypeS: sentencetypeR	-3.31080	1.47590	-2.243

The summary of the model in Table 3.3.2.2 shows that the main effect of Sentence Type is significant ($\beta = 2.680$, SE = 1.04 t = 2.571 for a sentence with a reflexive pronoun), suggesting that the other-oriented scenes elicit a higher amplitude for the reflexive pronoun than for the object pronoun. This is shown in the figures, too: the other-oriented pictures have a higher amplitude from 500-700 ms in Figure 3.3.2, which shows the reflexive pronoun, than in 3.3.1, with the object pronoun.

Moreover, it shows that the interaction effect between Picture Type and Sentence Type is significant ($\beta = -3.301$, SE = 1.476, t = -2.236 for a self-oriented picture with reflexive pronoun), indicating that a self-oriented picture cancels out the influence of Sentence Type. Figure 3.3.2 shows that the self-oriented conditions for the reflexive pronoun are indeed on the same level as the ones in Figure 3.3.2, for the object pronoun.

To summarize, the other-oriented, thus incongruent, pictures elicit a higher P600 in reflexive pronoun processing, indicating an effect of Picture Type. No effect of Introduction Order has been found on the P600.

Table 3.3.2.2: Fixed effects on P600 of the main effects model with interaction Picture Type x Sentence Type

Model: Cz ~ introtype + pictype + sentencetype + pictype : sentencetype + (1 ∣ Subject) + (1 ∣ Item)			
	Estimate	**Std. Error**	**t value**
(Intercept)	2.28750	0.98144	2.331
introtypeA2	0.89878	0.73027	1.231
pictypeS	-0.03159	1.05335	-0.030
sentencetypeR	2.68007	1.04238	2.571
pictypeS: sentencetypeR	-3.30079	1.47591	-2.236

3.4. Comparison between pupil dilation and EEG

Both pupil dilation and EEG have shown that Picture Type, Introduction Order, and the interaction between the two do not have an influence on the processing of an object pronoun.

A different effect between pupil dilation and brain activity can be seen in reflexive pronoun processing. Whereas in EEG, Picture Type has a main effect on reflexive pronoun processing (that is, other-oriented, thus incongruent, pictures elicit a higher P600), an interaction effect between Picture Type and Introduction Order occurs in pupil dilation (the other-oriented picture with actor-first introduction elicits a larger pupil dilation compared to the other conditions).

4. Discussion

Surprisingly, we did not find an effect of Picture Type or Introduction Order on object pronoun processing, whereas we did find an effect of Picture Type and Introduction Type on reflexive pronoun processing. This might have to do with the fact that in object pronoun processing, the listener knows that there are two antecedents to which the pronoun might refer. It is more ambiguous than the reflexive pronoun. Therefore, a picture that is incongruent or a non-canonical introduction order does not surprise the listener a lot. In reflexive pronoun processing, however, there is only one antecedent to which the reflexive can refer: the subject. In that case, if the scene is incongruent, the listener is more surprised. This would explain the fact that the incongruent pictures in EEG elicit a higher P600 than the congruent pictures. However, what still remains uncertain, is why in pupil dilation only the actor-first other-oriented condition in reflexive processing elicits a larger pupil dilation, whereas this interaction effect cannot be seen in EEG.

To answer the research question, which was: ‘Do visual and linguistic context (introduction order) have an influence on object pronoun processing in a blank screen paradigm? If there are effects:

- What are the effects on pupil dilation?
- Can these effects be seen in the EEG?
- When do they occur?’

we have to say that the current study has shown that visual and linguistic context do not have an influence on object pronoun processing as it cannot be seen in pupil dilation or in EEG. The influence of visual context can, however, be seen in EEG on the P600, and an interaction effect between visual and linguistic context can be seen in pupil dilation. The hypothesis stated that these later amplitude differences might reflect semantic and pragmatic integrative processes. In the current study, it appears to reflect a syntactic violation of expectancy.

In order to investigate these results and to see what caused the difference with earlier studies (e.g., van Rij, 2012; van Rij et al, 2016), more elaborate research needs to be done to look at earlier and later effects, for example in the introduction sentence or even after the offset of the test sentence. The fact that a non-canonical introduction order does not have an influence on pronoun processing (and reflexive processing in EEG) might be because the participants learned the task very quickly and did not pay attention to the introduction after a while anymore. They could have built up an expectation immediately after seeing the picture and only pay attention to the object/reflexive pronoun that followed in the test sentence. If this is not the case, we should see some effect of the introduction order directly after mentioning the first referent in the introduction sentence: if the non-actor is mentioned first, this should elicit a larger pupil dilation. If there is no difference in pupil dilation, this might explain why there is no influence of introduction order.

Moreover, the effect of Picture Type (and therefore congruency) might be seen later in pronoun processing: this would have to do with the fact that object pronouns are more ambiguous. Whereas reflexive pronouns can only refer back to one antecedent and it is thus immediately clear when the picture was incongruent, this might take some more time for the object pronoun. No difference has been observed between the object and reflexive pronouns in accuracy and reaction times, but maybe the conscious processing could have taken some more time, which would be reflected in a delayed pupil dilation, maybe even after the test sentence.

Some improvements could have been made in order to make the results clearer.

First of all, we made a rounding mistake for the first thirteen participants: instead of a fixed break of 200 ms between the introduction sentence and test sentence, this time was variable between 900 and 1300 ms. However, this did not influence the accuracy and/or reaction times. Therefore, we do not expect a significant influence on the results.

Moreover, some of the pictures that we used might have been ambiguous (e.g., is it a crocodile or a dinosaur?). However, during the instructions we told the participants that the names of animals and objects were not part of the experiment so this, too, did not influence the results significantly.

Besides that, other-oriented pictures were thought to be easier to process than self-oriented pictures. A comparison for the accuracy and reaction times between other-oriented pictures and self-oriented pictures has shown that there was no difference. Therefore, we do not expect this to have had an effect on the results.

Finally, we got some feedback that the experiment was quite dull and that people got distracted after a while. Since the accuracy did not decrease over time, we do not expect this to be a major influence.

All in all, we do not expect the improvements to have had a significant influence on the current study's results. However, the current study does not provide sufficient information to support or reject the influence of context on pronoun processing: for this, more elaborate research needs to be performed. Still, we found very surprising results on the processing of reflexive pronouns, which might set fire to much more research in this area!

5. Acknowledgements

First of all, I would like to thank Anouk Hoekstra for taking the time to record all of the sentences. Also, I would like to thank Robbert Prins and Petra van Berkum for drawing the pictures. Finally, a special thanks to Jacolien van Rij and Jelmer Borst for the supervision of and contribution to this project.

References

Altmann, G.T.M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. Cognition, 93(2), 79-87.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(4), 748-769.

Boersma, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer program]. Version 5.4.22. URL http://www.praat.org/.

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. Frontiers in Human Neuroscience, 7, 758.

Chomsky, N. (1981). Lectures on government and binding: the Pisa lectures. Dordrecht, The Netherlands: Foris Publications.

Chow, W.Y., Lewis, S., Phillips, C. (2014). Immediate sensitivity to structural constraints in pronoun resolution. Frontiers in Psychology, 5, Article 630.

Clifton, C., Kennison, S., & Albrecht, J. (1997). Reading the words her, his, him: Implications for parsing principles based on frequency and on structure. Journal of Memory and Language, 36(2), 276-292.

Delorme, A. & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics (pdf, 0.7 MB) Journal of Neuroscience Methods 134:9-21.

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. The QuarterlyJournal of Experimental Psychology, 63(4), 639-645.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. Journal of Memory and Language, 27(6), 699-717.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. Cognitive Science, 17(3), 311-347.

Hoeks, B., & Levelt, W. J. M. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. Behavior Research Methods, 25(1), 16-26.

Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. Language and Cognitive Processes, 23(5), 709-748.

Kennison, S. M. (2003). Comprehending the pronouns her, him, and his: Implications for theories of referential processing. Journal of Memory and Language, 49(3), 335-352.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. Journal of Psycholinguistic Research, 18(1), 5-19.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

SR Research Experiment Builder 2.1.140 [Computer software] (2017). Mississauga, Ontario, Canada: SR Research Ltd.

van Berkum, J. J. A., Koorneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. Brain Research, 1146, 158-171.

van Rij, J. (2012). Pronoun processing: computational, behavioral, and psychophysiological studies in children and adults. Chapter 6: Event-related pupil dilation tracks integration of contextual and grammatical information. Groningen. 115-132.

van Rij, J., Hollebrandse, B., & Hendriks, P. (2016). Children's eye gaze reveals their use of discourse context in object pronoun resolution. In A. Holler, & K. Suckow (Eds.), Empirical perspectives on anaphora resolution (pp. 267-293). (Linguistische Arbeiten; Vol. 563). Berlin/Boston: De Gruyter.

Zellin, M., Pannekamp, A., Toepel, U., & van der Meer, E. (2011). In the eye of the listener: Pupil dilation elucidates discourse processing. International Journal of Psychophysiology, 81(3), 133-141.