Vivian Schreier

2022

Integration of Restrictive and Non-Restrictive Relative Clauses

in

Robin Hörnig, Sophie von Wietersheim, Andreas Konietzko & Sam Featherston (eds.),

Proceedings of Linguistic Evidence 2020: Linguistic Theory Enriched by Experimental Data

pp. 479–498

Tübingen: University of Tübingen

https://publikationen.uni-tuebingen.de/xmlui/handle/10900/119301

In R. Hörnig, S. von Wietersheim, A. Konietzko & S. Featherston (eds.) (2022), *Proceedings of Linguistic Evidence 2020: Linguistic Theory Enriched by Experimental Data*. Tübingen: University of Tübingen. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/119301

Integration of Restrictive and Non-Restrictive Relative Clauses

Vivian Schreier[1]
University of Tübingen
vivian.schreier@uni-tuebingen.de

1 Introduction

Relative clauses (RCs) are some of the most commonly used structures in everyday language, nevertheless, their syntactic make-up and the different types are still topics for discussion. The most basic distinction divides the group of RCs into restrictive relative clauses (RRCs) as in (1a) and non-restrictive relative clauses (NRCs) as in (1b), which are often also called appositive relative clauses.

(1) a. Every man who loves his wife cleans the dishes.
b. John, who loves his wife, cleans the dishes.

Even though the two RCs look identical on the surface except for the comma placement, their function and their characteristics differ. Fundamentally, the RRC restricts the RC head noun and thus limits the set of referents under discussion, whereas the NRC provides additional information on an already sufficiently defined RC head.

The two types of RC however display a wide range of further differences, which have received significant attention in the literature (cf. Jackendoff, 1977; Emonds, 1979; De Vries, 2006; Arnold, 2007). Some of these differences even suggest differences of how and where RRCs and NRCs are integrated into the matrix clause. These for example include differences in intonation; an NRC is separated from its matrix clause through intonational breaks beforehand and afterwards, but an RRC is not intonationally marked in any way. An NRC can have its own focus-background structure, while an RRC is completely contained in the focus-background structure of the matrix clause. An RRC can only modify a nominal phrase, but an NRC is able to qualify various categories including any type of nominals, prepositional phrases (2a), adjectives (2b), verb phrases (2c), or even sentences (2d). In example (2) the RC head is always placed in square brackets.

(2) a. Kim put it [on his back], which was the right place. PP
b. Kim was [really nice], which I didn't think she would AP
c. Kim [won the race], which I didn't think she could VP
d. [Kim won the race], which was a relief. CP
(Arnold, 2007: 274)

The relative pronoun used to introduce an RRC can be *who*/*which*, *that*, or the ∅-pronoun, whereas the NRC can only be introduced by *who*/*which.* In terms of movement operations, the RRC is more flexible as it can be extraposed, while the NRC does not readily do so, as seen in (3).

(3) a. Some men appeared at the door that Mary had been insulting.
b. *These men appeared at the door, who Mary had been insulting.
(Emonds, 1979: 243)

[1] Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 75650358 – SFB 833

One final major difference concerns scope phenomena, for example in negative scope or binding. While a negative operator in the matrix clause extends into the RRC and negates the whole construction (4a), an NRC is exempt from the negation. An NRC thus also cannot contain Negative Polarity Items (4b) unless they are licensed by a negation placed in the NRC (4c).

(4) a. I didn't talk to a victim who saw {*something/anything} incriminating.
b. I didn't talk to the victim, who by the way saw {something/*anything} incriminating
c. I talked to the victim, who by the way didn't see {*something/anything} incriminating.
(adapted from Arnold, 2007: 277)

The second scope phenomenon is variable binding. Here it is often claimed in the literature that quantifiers may bind a variable contained in an RRC like in (5a) but are unable to establish a binding relation to a variable in an NRC like in (5b).

(5) a. *No plane* which has an engine in *its* tail is a failure.
b. **No plane,* which has an engine in *its* tail, is a failure.
(Arnold, 2007: 291)

Some controversial claims have been made in the literature and various conflicting examples have been presented as evidence for or against each claim. We will look at this controversy and the underlying data in greater detail in Section 2, as variable binding will serve as the test diagnostic in our own experiments presented in this article.

The question that is often discussed in relation to the contrast between RRCs and NRCs is whether the disparities in function and behaviour are based on a difference in form, namely whether the two types of RC are attached to the matrix clause at different positions. Over the years numerous syntactic analyses with potential attachments points have been proposed, a selection of which will be discussed in Section 2. So far, no conclusive answer has been reached, amongst other things also due to the insufficient or contradicting data used to support those models.

This paper aims to approach this old problem from a new perspective by presenting the results of two *Acceptability Judgement* experiments in Section 3. These will allow us to establish a sufficient and comparable set of data, on the basis of which we will analyse the adequateness of the models from the literature discussed in Section 2.

2 Models of RC Attachments

2.1 C*onstituency Hypothesis* vs *Orphanage Approach* for NRCs

Despite considerable work on the topic over some decades, the syntactic placement of RCs into their matrix clauses is still controversial. RRCs have widely received a unified account in which the RC is directly attached to the RC head noun.[2] This general structure has been illustrated in Figure 1, albeit some differences in labelling categories exist between them (Jackendoff, 1977; Emonds, 1979; Toribio, 1992; de Vries, 2006; Arnold, 2007).

[2] To my knowledge the only alternative is proposed by Thompson (1971), who assumes that RRCs are derived from a coordinated main clause.

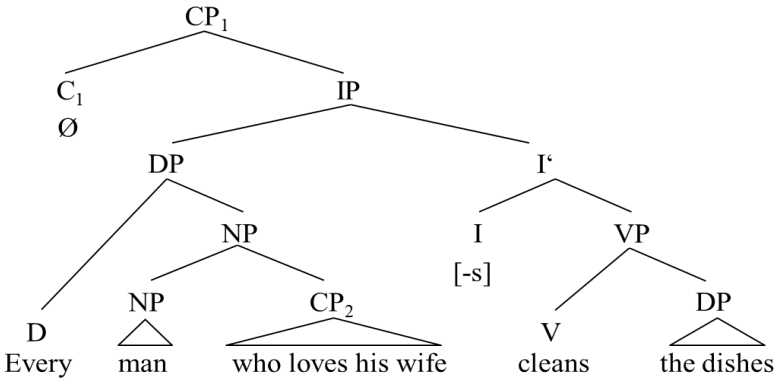

Figure 1. Restrictive relative clause

In Figure 1 the RRC (CP_2) is part of the matrix clause (CP_1) and adjoined directly at its head noun *man* as an adjunct. In contrast, the analyses for NRCs vary greatly. The *Constituency Hypothesis* (Smith, 1964; Jackendoff, 1977; Smits, 1989; Toribio, 1992; Platzack, 2000; de Vries, 2006; Arnold, 2007) assumes that the NRC and its head form one constituent, and thus adjoins the NRC within the matrix clause to the RC head. This is similar to the RRC analysis. However, within this hypothesis we find more fine-grained distinctions. I will focus on a strand which assumes right-adjunction (there are however alternatives which assume complementation, e.g. Smith, 1964; Platzack, 2000).

Even among the hypotheses which assume that the NRC is right-adjoined to its head subtle differences exist. Jackendoff (1977) and Smits (1989), for example, follow the NP theory of nominal phrases and assume a stack of N-layers. They postulate that the RRC attaches to lower N levels (N'' for Jackendoff, and N' for Smits), while the NRC is attached to the highest N level (N''' and NP respectively). Toribio (1992) on the other hand follows the DP theory of nominal phrases and considers the determiner as the head of a nominal. He analyses the RRC as attached to the NP and the NRC as attached to the DP, so that it is not within the scope of the determiner. A version of his NRC analysis can be seen in Figure 2. Here the NRC (CP_2) is also part of the matrix clause (CP_1) but serves as an adjunct to the DP *John*. Jackendoff (1977) and Smits (1989) propose the same structure but again would use different category labels.

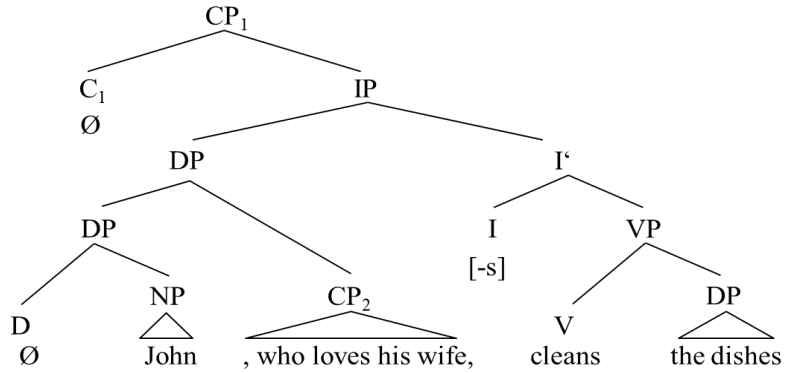

Figure 2. Non-restrictive relative clause attachment in the *Constituency Hypothesis* after Toribio (1992)

In contrast to the *Constituency Hypothesis*, a number of other analyses can be summarized under the term *Orphanage Approach*. One subcategory, the *Non-Radical Orphanage Approach* (Ross, 1967; Emonds, 1979; McCawley, 1982; Stuurmann, 1983), locates the NRC not at the nominal head of the RC but at the CP level, and employs movement operations in order to guarantee the linear adjacency of the RC head and the NRC. Emonds (1979) proposes that the NRC and its RC head end up in adjacent positions not by moving the RC but by extracting the intervening constituent and reattaching it to the right of the NRC, at the highest clause level.

This analysis has been discarded due to evidence from VP ellipsis (McCawley, 1982) and subjacency violations (Perzanowski, 1980). McCawley (1982) offers an alternative in which he also assumes that the base position of the NRC is at the CP level, however, he achieves the correct linearity of RC head and the NRC by introducing order-changing operations which do not affect the hierarchical structure of the constituents. In the hierarchical structure the NRC (CP_2) remains adjoined to the matrix clause CP_1, but in defiance of the general avoidance of discontinuous structures the NRC can then branch over other sentence constituents and insert itself next to its head in a sentence internal position. We see an illustration of this innovative approach in Figure 3.

Another subcategory, the *Radical Orphanage Approach* (Safir, 1986; Haegeman, 1988/2009;

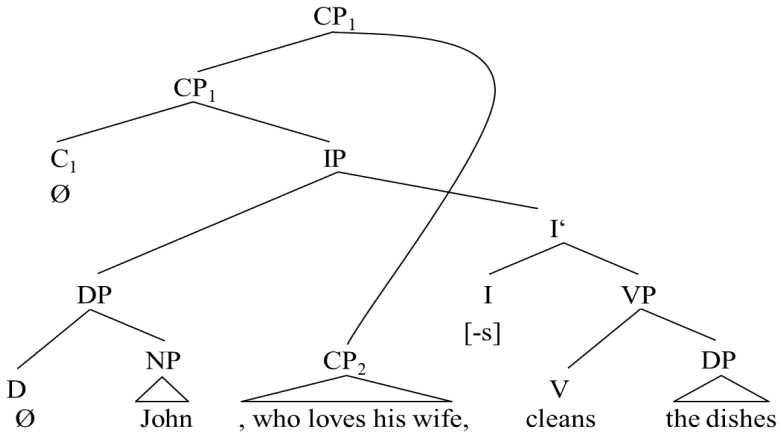

Figure 3. Non-restrictive relative clause in the *Non-Radical Orphanage Approach* after McCawley (1982)

Espinal, 1990; Fabb, 1990), negates any kind of structural relation between the NRC and the matrix clause. A possible example tree of this version can be seen in Figure 4.

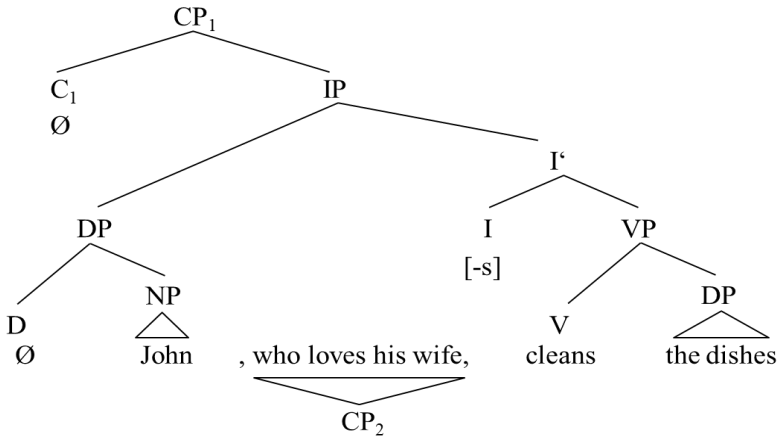

Figure 4. Non-restrictive relative clause in the *Radical Orphanage Approach* after Espinal (1990)

The NRC is considered an independent syntactic structure on a separate level or tier, which is only inserted next to the RC head at a post-syntactic stage. Safir (1986) assumes that this inclusion happens after LF at a level he then calls LF'; Espinal (1990) proposes that the various syntactic structures coexist in a three-dimensional space and only align on a phonetic level.

2.2 Previous Experimental Research

Even though these conflicting NRC analyses have been discussed at length, a consensual assessment has not been reached. One reason for this might be that the work done so far has been mainly theoretical, with little empirical validation. Experimental investigations could thus shed

new light on an old problem. The only previous attempt at an experimental investigation of RC attachment was Dillon et al. (2018), who investigate ambiguity resolution between RRC and NRC. They conducted a forced-choice experiment in which the participants had to answer which of the two potential referents the RC referred to in sentences like (6).

(6) a. *RRC_of*
Peggy ignored the child of the patient that had an annoying voice.
b. *RRC_with*
Peggy ignored the child with the patient that had an annoying voice.
c. *NRC_of*
Peggy ignored the child of the patient, who had an annoying voice.
d. *NRC_with*
Peggy ignored the child with the patient, who had an annoying voice.
(Dillon et al., 2018: 6)

Their premise for RRCs is that the different prepositions cause a different ambiguity resolution. The preposition *of* continues the thematic domain[3] of the object and thus the RC will be more often interpreted to modify the higher referent, in this case *the child*. By contrast, the preposition *with* creates a new thematic domain, which then serves as the most salient attachment site for the RC so that the RC is more often interpreted to modify the lower referent *the patient*. The *of*/*with* distinction forms a base line for the NRCs. Dillon et al. (2018) claim that if the NRCs display the same kind of ambiguity resolution pattern as the RRCs, this supports the *Constituency Hypothesis* and thus local attachment of both RRCs and NRCs. If, however, the NRCs are not locally attached, they should always favour the higher attachment site *the child*. The results confirm the distinction expected for the RRCs with more high attachments for the *of* conditions and more low attachments for the *with* conditions. The NRC cases show a similar pattern with a slight preference for low attachment, which Dillon et al. (2018) interpret as supporting the *Constituency Hypothesis*.

This is a very interesting approach and a first step to utilizing experimental data to distinguish between the different attachment models. Although the authors convincingly show a clear trend of the NRC being influenced by the *of*/*with* distinction, the question remains whether the *Constituency Hypothesis* offers the only explanation for this phenomenon. Dillon et al.'s (2018) premise that within the *Orphanage Approach* the NRC is expected to modify only the higher noun phrase *the child* is too strong because according to the different strands of the *Orphanage Approach*, the NRC is either attached to the CP layer or not at all. This means that the NRC would never be a syntactic part of the thematic domain of its RC head. Neither in the *of* condition where one thematic domain contains the two potential referents nor in the *with* condition where *the child* and *the patient* form two distinct thematic domains. However, the NRC linearly follows a thematic domain that either contains one or two potential referents, and in the absence of a syntactic connection to either of those referents the change of the thematic domain in the *with* condition might affect the ambiguity resolution. In this case the same pattern as for the RRC is expected although the syntactic structure underneath is not identical. The design of this experiment is thus not able to guarantee that the underlying syntactic attachment to either noun is responsible for the different ambiguity resolutions with the different prepositions.

2.3 Binding and C-Command as Diagnostic Tools for Integration

Our proposal will take a different experimental approach using variable binding as a diagnostic tool to observe differences of attachment between NRCs and RRCs. We speak of variable binding when a quantified noun phrase serves as an antecedent for the resolution of a pronoun.

[3] Dillon et al. define this as "the domain to which a non-primary item like the restrictive relative clause can be associated" (Dillon et al., 2018: 5).

Büring (2005) points out that this type of pronoun resolution is different from a normal coreference relation between a noun and a pronoun. In a coreference relation the pronoun is simply a substitution of a repetitious noun phrase, as can be seen in (7).

(7) a. *John* said that *he* was okay.
b. *John* said that *John* was okay.

(Büring, 2005: 81-82)

The same cannot be said when the pronoun antecedent is a quantified expression, here a simple substitution would not express the same meaning, in this case it would even lead to the contrary interpretation of the sentence.

(8) a. *No woman* doubts that *she* is okay.
b. *No woman* doubts that *no woman* is okay.

(Büring, 2005: 81-82)

Büring (2005) ascribes this difference to the fact that quantified expressions do not refer to any specific entity in the first place and any pronoun which should receive its interpretation from this quantified expression is thus unable to corefer. The anaphoric relationship between a quantified antecedent and its pronoun must thus be treated differently from a normal coreference relation and is subject to other restrictions.

These restrictions are explored first in detail by Reinhart (1983). Reinhart compares different restrictions which need to be fulfilled in order for a quantified expression or a *wh*-element to bind a pronoun. One such proposal of restrictions is based on precedence and command, meaning that the pronoun follows the quantifier and is its clause mate while neither dominates the other. Reinhart provides counterevidence to this through examples as in (9) where the pronoun precedes the quantifier and the sentence is still acceptable, or as in (10) in which the pronoun fulfils the criteria of following the quantifier and being its clause mate, but the sentence is unacceptable.

(9) Near *his* child's crib *nobody* would keep matches.

(Reinhart, 1983: 119)

(10) *People from *each of the small western cities* hate *it.*

(Reinhart, 1983: 124)

The fact that the restriction of the quantifier preceding and commanding the variable is not able to capture sentences like (9) and (10) leads Reinhart to propose an alternative hypothesis based on c-command, which she states as follows:

(11) Quantified NPs and wh-traces can have anaphoric relations only with pronouns in their c-command syntactic domain.

(Reinhart, 1983: 122)

(12) Node A c(onstituent)-commands node B iff the branching node α_1 most immediately dominating A either dominates B or is dominated by a node α_2 which dominates B, and α_2 is of the same category as α_1.

(Reinhart, 1983: 23)

C-command is given in example (9) if we assume that location *near his child's crib* has been preposed and its trace in its original position at the end of the sentence is still c-commanded by the quantifier *nobody*. In (10) c-command is not given because the quantified noun phrase is contained within the PP of the complex subject and thus only c-commands the preposition *from*. Reinhart's restriction on quantifier binding is thus able to explain the data set which poses problems for the precedence and command restriction.

The necessity for c-command in order to establish an anaphoric relation between the quantified noun phrase and the pronoun offers direct evidence of the underlying structure, which allows us to test the conflicting RC analyses using variable binding. In the most basic distinction of the *Constituency Hypothesis*, c-command should be given between the RC head and the RRC, as well as between the RC head and the NRC. However, in the *Orphanage Approach* the RC head should c-command the RRC, but not the NRC.

2.4 Problems with Binding

The situation is not as clear, however, because the intuitions reported in the literature about the acceptability of binding into different RC types are not consistent. As intuitions can be easily influenced by factors other than the phenomenon of interest, we must therefore be careful when drawing general conclusions from individual introspective judgements. It is thus important to check which other factors may have an influence on the overall acceptability of a sentence and how much of it can be attributed to a possible binding violation.

One such factor which potentially distorts the judgements can be seen in Arnold's (2007) observations of potential variable binding into NRCs. As a supporter of the *Constituency Hypothesis*, Arnold argues that a quantifier in the matrix clause can bind a pronoun in an NRC and that all contrary evidence is due to lack of plausibility. Arnold models his examples (13a-b) after Ross' (1967) examples and judgements of variable binding into RCs. Ross, as a supporter of the *Orphanage Approach,* judges the binding of a pronoun in an NRC as unacceptable. Arnold in turn acknowledges a difference in acceptability between the *a* and *b* sentences of example (13), however, he also claims that by making the context more plausible, as in the *c* version of the examples, the overall sentence and the binding become acceptable.

(13) a. *No plane* which has an engine in *its* tail is a failure.
b. **No plane,* which has an engine in *its* tail, is a failure.
c. *No modern plane*, which may or may not have an engine in *its* tail, is prone to this kind of problem.

(Arnold, 2007: 291-292)

Crucially, the adjustments made to the *c* sentence in order to increase its plausibility also significantly change the sentence. Specifically, Arnold includes the modal element *may or may not*, which allows for *Modal Subordination*, as Arnold himself notes in the last section of his paper. *Modal Subordination* is a phenomenon discussed by Roberts (1988) in which the use of modals or non-factual mood broadens the setting of an utterance in multiple possible worlds or common grounds, from which the discourse participants may choose the most appropriate. The consecutive, modally subordinated clauses or sentences then depend on the truth value of the first clause to fulfil their own truth values. The specifics of this phenomenon are of no particular interest here, however, it has been noted in the literature that *Modal Subordination* is able to make otherwise inaccessible binding relations, such as variable binding over clausal boarders, felicitous (Roberts, 1988; Arnold, 2007). It is thus unclear whether Arnold's acceptability rating of sentence (13c) is in fact due to general successful binding into an NRC or only due to *Modal Subordination* overruling a binding violation.

Nevertheless, there are various attested examples of successful variable binding into NRCs in the literature:

(14) *Every parrot* sang a song, which *it* didn't understand.

(Kempson, 2003: 302)

(15) *Every producer* paid the lead actress, who hates *his* guts, a fortune.

(Kamp & Reyle, 1993: 255)

(16) *Every rice-grower in Korea* owns a wooden cart, which *he* uses when *he* harvests the crop.

(Sells, 1985: 2)

It is remarkable that in all three examples, the quantified expression is the matrix subject, the NRC containing the pronoun however modifies the matrix object. This might point us to another factor which influences the acceptability of assumed binding violations of NRC sentences. Both Ross (1967) and Jackendoff (1977) claim that NRCs are unable to modify quantified expressions in general, i.e. independent of variable binding. They support this claim with sentences like (17).

(17) *Any man, who drives a Cadillac, is insane.

(Jackendoff, 1977: 175)

In contrast, Kempson (2003) claims that all non-negative quantifiers can serve as NRC heads as in (18), while Arnold (2007) even argues for negative quantifier heads as can be seen in (13).

(18) Each child, who the Head himself interviewed, said he was regularly bullied.

(Kempson, 2003: 302)

Admittedly, this poses a new set of problems. If an NRC modifying a quantified noun phrase is unacceptable in itself or if it is only able to modify a specific type of quantifier, judgements of unacceptability such as in (13b) cannot reveal anything about a potential binding violation because the asterisk marking might be caused by an independent factor. The diverging judgements found in the literature thus might be caused by a divergence of the source material, where other factors are causing the unacceptability. Sentences like (13b) might thus be unacceptable not because of a binding violation as claimed in the literature but because of the unacceptable combination of an NRC directly modifying a QP. We shall call this the *quantifier-head effect*.

Sentences like (14)-(16), where the binding antecedent is the quantified subject but the RC head is a non-quantified object, could thus on the one hand be acceptable because there really is no binding violation as claimed by the supporters of the *Orphanage Hypothesis*. On the other hand, in these sentences the type of quantifier might play a role in such a way that not every quantifier can serve as a head for an NRC.

This is an important issue: what we can take away from the judgements reported in the literature and the differences among them is that in order to gain reliable evidence it is important to control the factors that play into the judgements as strictly as possible. The type of quantifier, the placement of the RC, and surrounding structures and context can all affect the overall acceptability independently of binding behavior. The examples that are compared to gain the judgements, thus need to be as similar as possible in those regards.

In order to investigate these potential complications, we decided to conduct a preliminary experiment to shed more light upon these issues. The experiment is designed to investigate both the influence of *Modal Subordination* and the effect of different quantifiers as antecedents. Crucially, the experiment does not yet include any binding relations so that we are able to only look at potentially confounding phenomena. The experiment has a 3x4 design, the first factor being the type of relative clause consisting either of an RRC, an NRC, which was disambiguated with the particle *by the way,* or Arnold's version of an NRC containing *may or may not*. The second factor was the type of quantifier which we compared: *every, no, few, many.* The sentences were slightly adjusted to correctly inflect for the plural or singular verb requirement of the different quantifiers. An example item for the quantifier *every* can be seen in (19).

(19) a. *RRC*
Every nurse who works at a hospital hates the uniform.

b. *NRC_may*
   Every nurse, who may or may not work at a hospital, hates the uniform.
c. *NRC_btw*
   Every nurse, who by the way works at a hospital, hates the uniform.

The twelve conditions were applied to twelve lexicalisations and evenly distributed over 4 lists, which also contained 15 filler items. 40 participants took part via *Prolific*. The gathered judgements were normalized and can be seen in Figure 5 and 6.

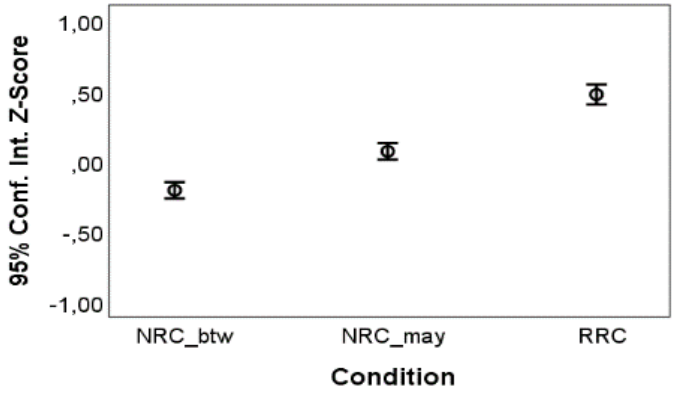

Figure 5. Results of the pilot study for different RCs

Turning first to the graph in Figure 5 we see the summarized results of the four different quantifier for the three different types of RCs we investigated in this experiment. The results show a clear preference for the RRC over the normal NRC (*NRC_btw*). The NRC containing *may or may not* is rated clearly worse than the RRC but still significantly better than the unaltered version of the NRC. We were thus able to show that the inclusion of modal constructions can affect the acceptability of a sentence. This means that we need to avoid such constructions in order to gain valid evidence on potential binding violation costs. However, the study also shows that the NRC condition containing *by the way* is significantly worse than the RRC condition even though we did not include any anaphora resolution. This might either mean that NRCs are generally less acceptable than RRCs, or that NRCs modifying quantified expressions are less acceptable than RRCs modifying the same heads. We will have to take this into account too when interpreting our further studies in order to obtain fully controlled evidence.

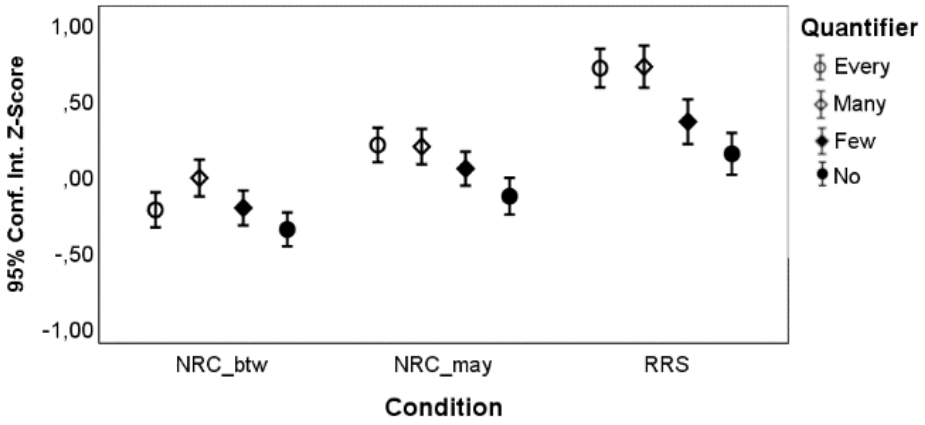

Figure 6. Results of the pilot study split up for different quantifier

The graph in Figure 6 gives a detailed picture of how the different quantifiers reacted to the different RCs. The findings of this study showed that all tested quantifiers can serve as heads for RRCs and NRCs because they follow the same pattern for all three different versions of RCs. The quantifier *every* and *many* are mostly on the same level and always the most acceptable, while *few* is always slightly less acceptable and no is again slightly less acceptable than *no*.

We decided to employ the universal quantifier *every* in our binding experiment because it provides the most discernible and reliable results and picks out a singular complement, which will play a crucial role in the ambiguity resolution in our next experiments. Furthermore, *every* had also been the quantifier most commonly used in the attested examples in the literature (14) - (16).

After identifying some of the factors which may have led to contradictory introspective judgements in the literature, we then conducted a number of judgement experiments to create a more reliable data set from which to draw our conclusions about which of the two major strands of analyses, *Constituency* or *Orphanage*, can capture the integration behaviour of RCs. The strictly controlled material as well as the procedure to gain the judgements will be illustrated in the next section.

3 Acceptability Judgement Studies

3.1 Experiment 1

3.1.1 Material and Methodology

The data presented here stems from our first experiment with three binary factors.

The first factor is BINDER, which contrasts variable binding (QP antecedent) with co-reference pronoun resolution (DP antecedent). This aims to compare the results drawn from variable binding, which depends on c-command, to normal pronoun resolution, which does not. We chose *every* as the quantifier for the QP conditions and alternated between *a* for RRCs and *the* for NRCs for the DP antecedent to meet the requirements of specificity of the different RC types.

The second factor, RC HEAD, is based on another finding from our pre-experiment, namely, that even though it is not completely impossible for NRCs to modify a QP, it is consistently less acceptable than for RRCs. As this coincides with our observation from the attested examples in the literature, where mainly sentences with the RC modifying the unquantified object received acceptable judgements, we decided to make this a factor in our design to check for biases. We thus alternated between attaching the RCs to the matrix subject and the matrix object. The matrix subject c-commands the matrix object and all its constituents. Note that the object was always a DP in order to establish whether different judgements found in the literature for binding behaviour of RRCs and NRCs are caused by this problem. The possessive pronoun that should be potentially bound in the RRCs or NRCs could only get its interpretation from the matrix subject: matrix subject and potentially bound pronoun were matched in gender, whereas matrix object and pronoun were mismatched.

As the last factor, RC TYPE, we compared RRCs and NRCs. We guaranteed the different interpretations between those two forms by including the necessary commas for the NRCs as well as including the discourse particle *by the way*. An RC containing a discourse particle such as *by the way* – which points out additional information – can only be construed as non-restrictive and thus allows us to eliminate potential misinterpretation of the sentences (Busch & Schumann, 2016).

Example (20) shows all conditions in one of the lexicalisations from our material.

(20) a. *QP_S_RRC*
Every landlord who looks after his property employs a cleaning lady.
b. *QP_S_NRC*
Every landlord, who by the way looks after his property, employs a cleaning lady.
c. *DP_S_RRC*
A landlord who looks after his property employs a cleaning lady.
d. *DP_S_NRC*
The landlord, who by the way looks after his property, employs a cleaning lady.
e. *QP_O_RRC*
Every landlord employs a cleaning lady who looks after his property.
f. *QP_O_NRC*
Every landlord employs the cleaning lady, who by the way looks after his property.
g. *DP_O_RRC*
A landlord employs a cleaning lady who looks after his property.
h. *DP_O_NRC*
The landlord employs the cleaning lady, who by the way looks after his property.

The label for each condition is made up by the sequence of the three factors, the first being BINDER, so either DP or QP, the second factor RC HEAD is either S for subject modification or O for object modification and the last factor RC TYPE is either RRC or NRC.

These eight conditions were realized in 16 lexicalisations, which were kept as similar in length, complexity, and plausibility as possible. The resulting 128 experimental stimuli were distributed over 16 counterbalanced lists and randomized with 15 standardized filler items (Gerbrich et al., 2019: 315). The participants were presented with one stimulus at a time and asked to judge its naturalness in comparison to two anchored reference sentences according to the *Thermometer Judgement* method (Featherston, 2008). We collected the judgements of 80 native speakers using the *Prolific* informant portal.

The different integration hypotheses result in contrasting predictions for the outcome. The *Constituency* school would basically predict no difference between any of the conditions because the structure would allow for the matrix subject to c-command into both an RRC and an NRC so that the binding relations should be equally successful. The *Orphanage Approach* on the other hand would predict that the NRCs are significantly worse than the RRCs because of the missing c-command relation between the quantifier and the pronoun in the NRCs, which leads to a binding violation. The other two factors give us opportunity to check for additional influences on the acceptability of our conditions. Because co-reference is not dependent on c-command to establish anaphora interpretation, the DP-antecedent conditions should not show any difference between the RRCs and the NRCs. If they nevertheless do, this might hint at one of them being generally harder to process than the other. The subject-object distinction checks for the potential *quantifier-head effect*, which should only apply to the subject conditions. The comparison of the subject and object RC-head conditions allows us to quantify this effect. Fundamentally, the object conditions yield the more reliable evidence. Taking all this into account, in order to support the *Orphanage Hypothesis,* we should find a significant interaction for BINDER and RC TYPE for both the subject and the object conditions.

3.1.2 Results

The judgements gathered were normalised into z-scores by subtracting the informants mean value from each judgement and dividing it by the person's standard deviation. This allows us to observe the results with a smaller degree of inter-informant variation.

In Figure 7, we see the error bars for both variable binding and co-reference for the two RC types. RRCs are always in white and NRCs in black. On the x-axis we see the structural conditions, with the first four error bars displaying the RCs attached to the matrix subject and the next four error bars having the RCs modify the matrix object. The QP conditions with variable binding always precede the DP conditions with co-reference. To the right-hand end of the graph we have the five error bars for the cardinal well-formedness values derived from the standard filler sentences (Gerbrich et al., 2019). They each consist of the mean of three standard items and give a point of reference for the overall acceptability of our experimental conditions.

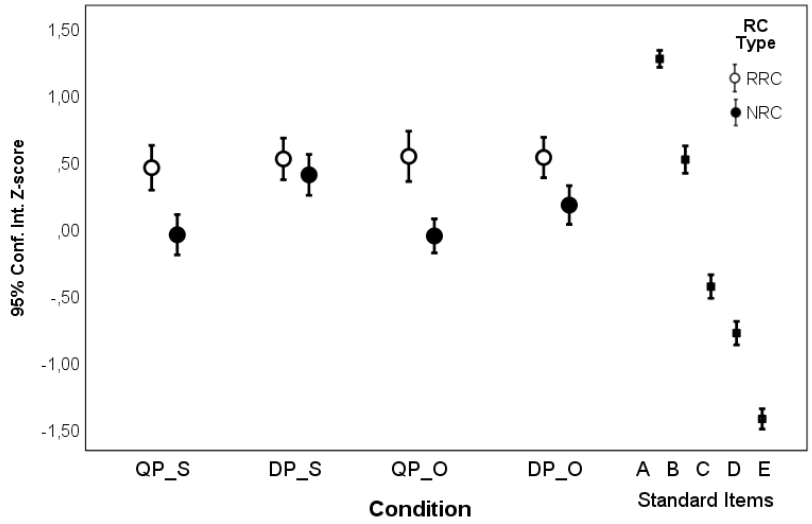

Figure 7. Results of the first experiment of the acceptability judgement study

All four error bars for the RRCs are all on the same level at .5, about equivalent to our B well-formedness value. The NRCs display more variation in their acceptability. Only the DP_S_NRC condition (NRC modifies subject, which in turn serves as an antecedent for the coreference relation of the pronoun) is as good as the RRC conditions, the DP_O_NRC (NRC modifies object, DP subject co-refers with pronoun) is slightly worse and the two QP conditions are even worse at roughly. However, all NRC conditions are still at or above the C level of the cardinal well-formedness values, which is made up of sentences of middling acceptability.

We ran a multifactorial ANOVA consisting of our three factors BINDER, RC HEAD, and RC TYPE. Over the whole data set we only got one significant interaction between BINDER (QP or DP) and RC TYPE (NRC or RRC): $F_1(79,1) = 8.802$; $p_1 < .005$; $F_2(15,1) = 5.897$; $p_2 < .03$; and two significant main effects for the two factors involved therein, (BINDER: $F_1(79,1) = 12.280$; $p_1 < .001$; $F_2(15,1) = 18.189$; $p_2 < .001$; RC TYPE: $F_1(79,1) = 35.566$; $p_1 < .001$; $F_2(15,1) = 28.858$; $p_2 < .001$).

Due to the concerns raised by the conflicting judgements in the literature, we also ran separate ANOVAs for only the subject conditions and only the object conditions. Those had only the two remaining factors, BINDER and RC TYPE. For the four conditions in which the subject serves as the head for the RC (20a-d) the ANOVA showed nearly the same significant effects as before. The interaction between BINDER and RC TYPE was still significant by subjects ($F_1(79,1) = 8.140$; $p_1 < 0.006$), but not by items ($F_2(15,1) = 3.7156$; $p_2 < .073$). The two main effects remained significant (BINDER: $F_1(79,1) = 5.296$; $p_1 < .001$; $F_2(15,1) = 24.415$; $p_2 < .001$; RC TYPE: $F_1(79,1) = 7.714$; $p_1 < .001$; $F_2(15,1) = 12.923$; $p_2 < .003$). For the four object conditions (20e-h) the ANOVA only resulted in one significant main effect for RC TYPE: $F_1(79,1) = 41.115$; $p_1 < .001$; $F_2(15,1) = 2.286$; $p_2 < .001$.

3.1.3 Discussion

The results drawn from the judgements study do not fully comply with the predictions of either theory and thus need to be evaluated very carefully.

Turning first to the RRCs, here the result is very clear and corresponds to what both hypotheses predicted. All conditions end up equally acceptable, meaning that neither the type of binding antecedent (DP or QP) nor the placement of the RC (modifying the matrix subject or object) have an effect on the overall acceptability of RRCs. This allows us to draw two important conclusions.

The first conclusion is that successful variable binding is as acceptable as co-reference. Even though this has never been discussed in the literature, it could have been a possible that properties of quantifiers, like non-specificity and non-referentiality (Reinhart, 1983; Büring, 2005), could have led to lower ratings of our QP conditions with variable binding compared to our DP conditions with co-reference. However, this was not the case: the co-reference conditions can truly serve as control conditions because both versions of anaphora resolution are equally acceptable. Linked to this is an additional issue related to this first conclusion, namely that the analysis of RRCs as being c-commanded by the subject, which both the *Constituency Hypothesis* and the *Orphanage Approaches* assumed, seems to be correct. If Thompson's (1971) right-adjunction analysis for RRCs without c-command was correct, we would also expect to see their QP conditions to be less acceptable.

The second conclusion pertains the RC head. The RRC can modify both DPs and QPs equally well, as the literature has suggested. Otherwise we should have found a difference in acceptability between the two subject conditions which would not show up for the object. The fact that absolutely no differences occur makes the RRCs an ideal baseline against which to contrast the behaviour of the NRCs.

For the NRC conditions the situation is a bit more difficult and the results paint a rather mixed picture, which does not fully confirm either of the two stands of analyses argued for by the *Constituency* or *Orphanage* supporters.

Since the *Constituency Hypothesis* would predict the same result for RRCs and NRCs, our statistical analysis set out to test the *Orphanage Approach* as the alternative hypothesis. To support the lack of integration and therefore of c-command the ANOVA must show a significant interaction between the factors RC TYPE and BINDER, meaning that the difference between the NRCs and the RRCs is significantly different in the QP conditions, but not in the DP conditions. Based on the literature we assumed that the NRCs are worse than the RRCs. Over all conditions this interaction is significant. From this alone we would be able to conclude that for the co-reference condition, which is not dependent on c-command, RRCs and NRCs are of equal acceptability, while for the variable binding conditions, which need c-command to establish felicitous anaphoric relations, NRCs are less acceptable. This would be in line with the *Orphanage Hypothesis* of the NRC either being too high up in the tree or completely external to the matrix tree to be c-commanded by the subject.

If we stopped our data exploration at this point, our results would show clear support for the *Orphanage* analysis. However, some authors have claimed that NRCs are less able or not able to modify a quantified expression (Ross, 1967; Jackendoff, 1977). This *quantifier-head effect* can cause an NRC sentence to be judged as less acceptable, which could result in the erroneous conclusion that this is brought upon by a binding violation. To control for this bias, we decided to compare RCs attached to the matrix subject and matrix object. In both cases the quantified subject is the binding antecedent from which the possessive pronoun receives its interpretation, but in the object condition the NRC does not modify this subject (cf. 16f). We therefore also ran two separate ANOVAs for the factor RC HEAD. For the conditions where the RC is attached to the subject, we got a similar result as described before with a significant interaction between RC TYPE and BINDER. However, this outcome could be caused by either a

real binding violation through the lack of c-command in NRCs, or the *quantifier-head effect*. In the light of this, we must regard the object condition results as more reliable because the *quantifier-head effect* does not apply here. Thus if we find an interaction in these conditions, it can only be caused by a binding violation. As reported in the results, the object conditions do not show a significant interaction. Based on these findings, we could conclude that the interaction which we find only for the subject conditions and which influences the outcome sufficiently also to show up when we run an ANOVA over all conditions is caused by the *quantifier-head effect* and not by an unsuccessfully bound pronoun. From this viewpoint, our experiment would provide evidence for the *Constituency* theory because no binding violation for the NRC sentences with QP subjects would mean that the underlying structure allows for the subject to c-command the NRC, and thus the NRC is fully integrated.

However, also this line of argumentation is not fully capable of explaining the distribution of data as we found it in our study. Turning back to results from the graph repeated here in Figure 8 we see that the two NRC error bars for subject and object attachment with variable binding are equally low.

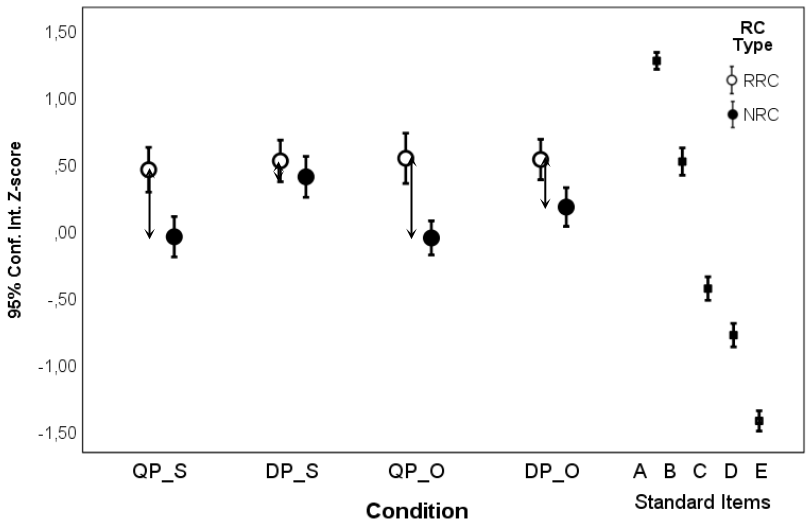

Figure 8. Results of the first experiment of the acceptability judgement study (rep)

This is contradictory to the predictions of the *Constituency Hypothesis* even with the added factor of the *quantifier-head effect* because the prediction would then be that the condition QP_O_NRC (17f) is as good as QP_O_RRC (17e). This is clearly not the case. What we find instead is that the behaviour of the QP conditions is the same for subject and object, meaning that the RRC error bars are at a level with the B standard items but the NRC error bars are very close to the C level. Different behaviour between subject and object attachment can be seen for the DP conditions because while RRCs and NRCs are at the same level for the conditions with the subject as the head, the NRCs are less acceptable than the RRCs for the conditions with the object as the head. This is the cause of the RC TYPE main effect which showed up in the object specific ANOVA. The marked acceptability of the DP_O_NRC condition and the RC TYPE main effect is unexpected for both accounts of analysis. The co-reference conditions do not rely on c-command to establish the anaphoric relation to the pronoun thus neither hypothesis would predict any difference for the DP conditions regardless of whether the RC modifies the subject or the object. A potential explanation that NRCs are generally harder to process since they introduce new propositions is contradicted by the DP_S_NRC error bar, which is not affected by any NRC dispreference. We must thus conclude that an additional factor has a negative impact on the DP_O_NRC condition compared to the DP_O_RRC condition. Even though this factor is not related to variable binding and c-command, the conclusion which analysis is supported by our experimental evidence remains inconclusive until we can comprehend this factor

better. For this purpose, we conducted a second experiment to investigate the impact of this unknown factor in more detail.

3.2 Experiment 2

3.2.1 Material and Methodology

Experiment 2 is a full repetition of Experiment 1 with additional conditions. We kept our initial three factors, BINDER, RC HEAD, and RC TYPE, to test the reproducibility of our previous findings. To clear up the remaining questions about a potential dispreference for the NRCs or an additional effect of definiteness on our DP_O conditions we introduced a new factor PRONOUN. Here we compared the sentence acceptability of the conditions containing a pronoun which needs an anaphoric relation as in (17) with identical conditions without a pronoun. In the latter conditions the possessive pronoun within the RC has been replaced with the definite article *the*. Without the pronoun there is no need for variable binding and thus no binding violations. However, all other influences such as the *quantifier-head effect* or additional processing efforts for NRC conditions should still show up. The set of new conditions thus serve as a baseline of the acceptability of our material against which to measure the potential additional binding violation cost of unsuccessful variable binding. The eight new conditions are listed in (21).

(21) a. *QP_S_RRC_noPrn*
Every landlord who looks after the property employs a cleaning lady.
b. *QP_S_NRC_noPrn*
Every landlord, who by the way looks after the property, employs a cleaning lady.
c. *DP_S_RRC_noPrn*
A landlord who looks after the property employs a cleaning lady.
d. *DP_S_NRC_noPrn*
The landlord, who by the way looks after the property, employs a cleaning lady.
e. *QP_O_RRC_noPrn*
Every landlord employs a cleaning lady who looks after the property.
f. *QP_O_NRC_noPrn*
Every landlord employs the cleaning lady, who by the way looks after the property.
g. *DP_O_RRC_noPrn*
A landlord employs a cleaning lady who looks after the property.
h. *DP_O_NRC_noPrn*
The landlord employs the cleaning lady, who by the way looks after the property.

The labels remain the same as in Experiment 1; the only new addition for the new conditions is *noPrn* for the conditions without a pronoun.

Together with the conditions in (20) we now have 16 conditions in eight lexicalisations, which were distributed over eight lists and mixed with the same set of standard items as in the previous experiment.[4] All other procedures and tasks remained identical and we collected the judgements of 48 participants who had not participated in Experiment 1via *Prolific.*

Through the additional factor PRONOUN the prediction to support the *Orphanage Hypothesis* is now a two-way interaction between BINDER, RC TYPE, and PRONOUN for both the subject and the object attachment.

3.2.2 Results

Figure 9 shows the normalized judgements as error bars. The structural conditions are placed on the x-axis, the first eight error bars are the identical conditions tested as in Experiment 1, the following eight error bars with the additional label *noPrn* are the new conditions without a

[4] We only used eight lexicalisations from the previous experiment and only distributed the items over eight lists to keep the experiment shorter.

possessive pronoun as illustrated in (21). At the right end of the graph are the five error bars for the standard levels of acceptability.

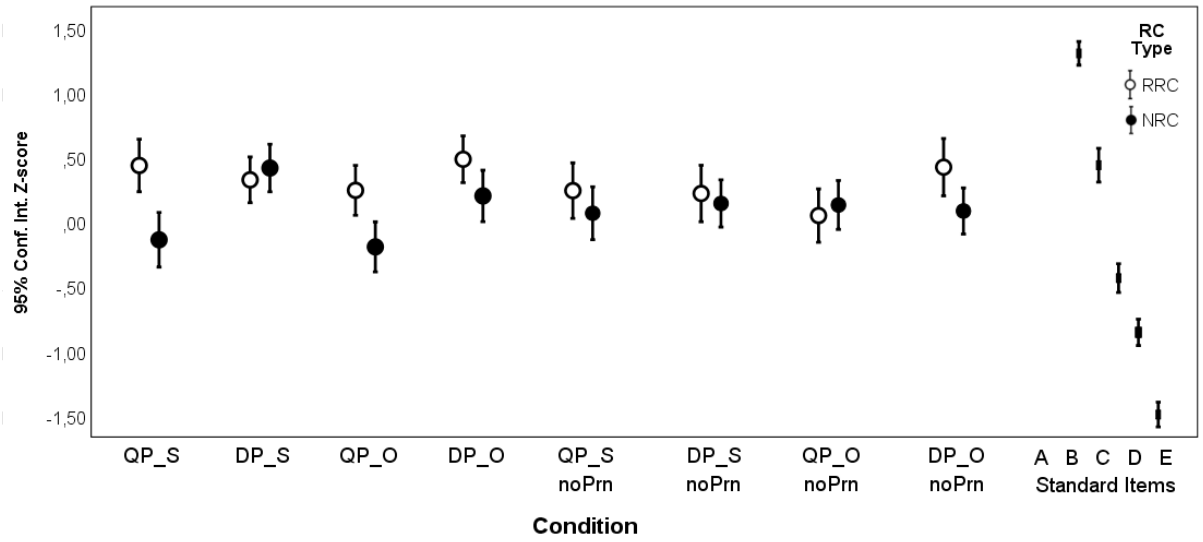

Figure 9. Results of the second experiment of the acceptability judgement study

The first ANOVA we ran was across all four factors. Here we found a significant two-way interaction between BINDER, RC TYPE, and PRONOUN: $F_1(47,1) = 1.463$; $p_1 < .002$; $F_2(7,1) = 13.352$; $p_2 < .008$, as well a main effect for BINDER ($F_1(47,1) = 9.041$; $p_1 < .004$; $F_2(7,1) = 13.082$; $p_2 < .009$) and a main effect for RC TYPE in the analysis by subjects ($F_1(47,1) = 9.660$; $p_1 < .003$; $F_2(7,1) = 5.488$; $p_2 < .052$). On top of that we also had significant interactions in the analysis by items between RC TYPE and PRONOUN ($F_1(47,1) = 3.113$; $p_1 < .084$; $F_2(7,1) = 8.165$; $p_2 < .024$,) and BINDER and PRONOUN ($F_1(47,1) = 3.473$; $p_1 < .069$; $F_2(7,1) = 14.356$; $p_2 < .007$).

Based on the interactions we found and motivated by the literature and our first experiment we again ran separate analyses for the factor RC HEAD. For RC attached at the subject the interaction between BINDER, RC TYPE, and PRONOUN was still significant in the analysis by subjects but not by items: $F_1(47,1) = 5.891$; $p_1 < .019$; $F_2(7,1) = 4.437$; $p_2 < .073$. The simple interaction between BINDER and RC TYPE was also significant ($F_1(47,1) = 5.462$; $p_1 < .024$; $F_2(7,1) = 1.506$; $p_2 < .014$) as well as a RC TYPE main effect for the analysis by subjects ($F_1(47,1) = 4.419$; $p_1 < .041$; $F_2(7,1) = 2.741$; $p_2 < .142$.) The separate analysis for the conditions in which the RC modifies the object also show a significant interaction between BINDER, RC TYPE, and PRONOUN: $F_1(47,1) = 6.623$; $p_1 < .037$; $F_2(7,1) = 8.666$; $p_2 < .022$ as well as a main effect for BINDER ($F_1(47,1) = 12.096$; $p_1 < .001$; $F_2(7,1) = 6.065$; $p_2 < .043$) and a main effect for RC TYPE in the analysis by subjects ($F_1(47,1) = 6.373$; $p_1 < .015$; $F_2(7,1) = 3.372$; $p_2 < .109$).

In a final step we also ran separate analysis for the new factor PRONOUN. For conditions with a pronoun and the RC modifying the subject, which are the conditions also tested in Experiment 1, the analysis yielded a significant interaction between BINDER and RC TYPE ($F_1(47,1) = 13.726$; $p_1 < .001$; $F_2(7,1) = 15.100$; $p_2 < .006$) and a main effect for BINDER ($F_1(47,1) = 6.218$; $p_1 < .016$; $F_2(7,1) = 16.836$; $p_2 < .005$). The parallel conditions without a pronoun had no significant results. For the set of conditions in which the object is the head for the RC containing a pronoun, which was again equivalent to the conditions from the first experiment, the results showed significant main effects for BINDER ($F_1(47,1) = 12.203$; $p_1 < .001$; $F_2(7,1) = 2.926$; $p_2 < .003$) and RC TYPE ($F_1(47,1) = 14.141$; $p_1 < .001$; $F_2(7,1) = 8.834$; $p_2 < .021$), but no interaction. The parallel new conditions for *noPrn* show no main effects but an

interaction between BINDER and RC TYPE ($F_1(47,1) = 5.684$; $p_1 < .021$; $F_2(7,1) = 9.806$; $p_2 < .017$).

Figure 10 summarizes the results from the last ANOVA by indicating which comparisons yielded significant interactions or main effects.

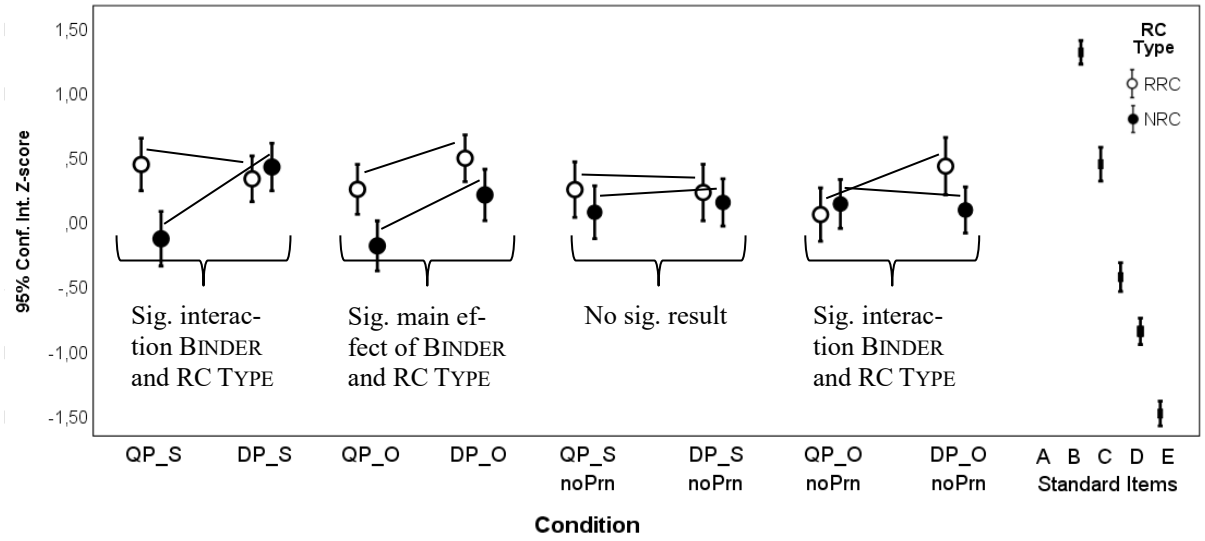

Figure 10. Significant interactions and main effects of the second experiment of the acceptability judgement study

3.2.3 Discussion

Turning first to the eight new conditions without a pronoun, we see that with the exception of the DP_O conditions we find no difference between the RRC and NRC error bars. For the subject conditions we found no statistical effects, whereas for the object conditions we found an interaction between BINDER and RC TYPE, meaning that there is more difference between DP_O conditions than between the QP_O conditions as Figure 10 shows. This difference between the RRCs and the NRCs in the DP_O_noPrn condition is the exact same difference that we had already found in the DP_O condition in our first experiment. We can thus conclude that the unknown effect is stable and based in our basic sentence material; it is not caused by any effect of binding or co-reference relations. The lack of a difference between the RRCs and the NRCs for both QP_noPrn conditions and the DP_S_noPrn shows that that there is no general preference of RRCs over NRCs. This leads us to conclude that an additional factor applies to the DP_O conditions, which does not affect the other conditions. It is thus necessary to look at the material for this condition once again. The two sentences in question are repeated here in (22) and (23).

(22) a. *DP_O_RRC*
A landlord employs a cleaning lady who looks after his property.
b. *DP_O_NRC*
The landlord employs the cleaning lady, who by the way looks after his property.

(23) a. *DP_O_RRC_noPrn*
A landlord employs a cleaning lady who looks after the property.
b. *DP_O_NRC_noPrn*
The landlord employs the cleaning lady, who by the way looks after the property.

The difference between the *a* and *b* sentences of (22) and (23) is the use of articles. In order to fulfil the requirements for RRCs to modify an indefinite head and NRCs to modify a definite head while keeping the material as parallel as possible to their respective DP_S conditions the RRC conditions have both an indefinite subject and object, while the NRC sentences have two

definite referents. We assume that the NRC sentences were thus rated worse because the use of the definite article implies discourse-givenness compared to an indefinite article, which is used when an entity is first introduced. As we did not use any context sentences all referents were discourse-new and the repetitious use of the definite article in our DP_O conditions might have appeared unnatural. Interestingly, the corresponding DP_S sentences, where only the subjects differed in definiteness (indefinite subject for RRC, definite subject for NRC), while the objects were indefinite for both sentences, were not significantly affected by this.

Having established that the difference between the RRCs and NRCs in the DP_O condition can also be found in the DP_O_noPrn condition, we are now able to explain the overall results in the original experiment and their repetition in Experiment 2 better. The eight conditions which were a repetition of Experiment 1 had to a large extent the same significant effects as was shown in the separate ANOVAs for the PRONOUN factor. The subject conditions have the same significant interaction between BINDER and RC TYPE as in Experiment 1, which could either be caused by a binding violation or the *quantifier-head effect* as we discussed before. The object condition had again a main effect for RC TYPE and a new main effect for BINDER. Overall, this shows that Experiment 2 was successful in replicating Experiment 1 and that all measured effects were stable when repeated. We must thus conclude that our initial prediction that the results should show an interaction between RC TYPE and BINDER for both subject and object attachment were wrong. This prediction is correct for the subject conditions and also finds support in our data but the object conditions are more complicated. We see that the difference in the material when it comes to the definiteness of the articles used already creates a negative bias for the NRC sentences for all DP_O conditions. This shows up as interaction between BINDER and RC TYPE in the base material. It is important to note that this interaction has an inverse direction to what we predicted for the conditions containing a pronoun; the difference between RRCs and NRCs is bigger for the DP_O_noPrn condition than for the QP_O_noPrn condition. The RC TYPE main effect we observed in both Experiment 1 and 2 for the object conditions is thus caused by the added effect of the interaction given in the base material and the predicted interaction through the violation cost of a quantifier serving as an antecedent of a pronoun without c-commanding it. The violation cost pushes the QP_O_NRC condition down to the same level as the DP_O_NRC and thus levels out the interaction in the base material resulting only in a RC TYPE main effect.

This finding is further supported by the overall analysis of Experiment 2. In our predictions we stated that a two-way interaction between BINDER, RC TYPE, and PRONOUN for both the subject and the object attachment was necessary to support the *Orphanage Approach*. This two-way interaction was found both over all conditions as well as in the two separate ANOVAs for the subject and object conditions. We can thus conclude that both Experiment 2 as well as Experiment 1, when we discount the additional effect on the DP_O conditions, show support for the *Orphanage Approach*. In a broader sense our results might also be able to explain how the contradictory judgements in the literature come about: even though the QP_NRC conditions are significantly worse than their RRC counterparts, they are still on one level with the C cardinal well-formedness value of the standard items. That means that on an absolute scale they are still of middling acceptability, which makes it understandably that some linguists would still consider them as acceptable, while others would consider them marked.

4 Conclusion

Our study aimed to compare different models of NRC attachment, which have been proposed in the literature and test the differing claims about potential variable binding in acceptability judgement experiments. While the *Constituency Hypothesis* assumes that the NRC is attached locally at its head and thus enters into a c-command relation with the subject of the matrix clause and allows for successful binding of a pronoun within the RC, the *Orphanage Approach* locates the NRC either at the highest syntactic level of the clause or completely outside the syntactic structure. This in turn then negates the presence of c-command between the subject

and the entire NRC, and should thus also block a quantified subject from successfully binding a pronoun within the RC. Since the literature reported conflicting introspective judgements regarding this phenomenon which were often influenced, as we were able to show, by additional factors such as *Modal Subordination* or the acceptability of a quantified expression to serve as a head for both RC types, we conducted two rating experiments to compare the acceptability of QPs and DPs entering in an anaphoric relation with a possessive pronoun contained in an NRC or RRC. Our results of both experiments show clear support for the *Orphanage* analysis. For both the subject and the object conditions we find reduced acceptability of the NRC compared the RRC. The predicted interaction between the RC TYPE and the BINDER is statistically significant for the subject conditions, whereas the main effect RC TYPE shows evidence of the same effect for the object conditions once allowance is made for the pre-existing difference in the base material due to the use of definite articles. Since variable binding is regarded as a syntactic process, which is based upon a c-command relation between the quantifier in the matrix clause and the variable placed in the RC, this implies that NRCs are located in a position which is not c-commanded by the subject of the matrix clause. These findings are consistent with those structural hypotheses which claim that NRCs are not integrated into their head NPs, but rather are attached at a clausal level, or are not syntactically integrated at all (McCawley, 1988; Espinal, 1991; contra Jackendoff, 1977; Arnold, 2007). Unfortunately, a more fine-grained distinction between the different strands contained within the *Orphanage Approaches*, such as the *Radical* or *Non-Radical Orphanage Approach* is outside the interpretive scope of our findings.

References

Arnold, D. (2007). Non-restrictive relatives are not orphans. *Journal of Linguistics, 43,* 272-309.

Büring, D. (2005). *Binding Theory*. Cambridge: Cambridge University Press.

Busch, J., & Schumann, F. (2016). Unspecific indefinites and (non-)restrictive relative clauses. *Lingua, 181*, 1-35.

de Vries, M. (2006). The syntax of appositive relativization: On specifying coordination, false free relatives, and promotion. *Linguistic Inquiry, 37*(2), 229-27.

Dillon, B. Frazier, L. & Clifton, C. (2018). No longer an orphan: evidence for appositive attachment from sentence comprehension. *Glossa, 3*(1)*:32*, 1-22.

Emonds, J. (1979). Appositive relatives have no properties. *Linguistic Inquiry, 10*(2), 211-243.

Espinal, M. T. (1991). The representation of disjunct constituents. *Language*, *67*(4), 726-762.

Fabb, N. (1990). The difference between English restrictive and nonrestrictive relative clauses. *Journal of Linguistics, 26*(1), 57–77.

Featherston, S. (2008). Thermometer judgments as linguistic evidence. In C. M. Riehl & A. Rothe (Eds.), *Was ist linguistische Evidenz?* (pp. 69-90). Aachen: Shaker Verlag.

Gerbrich, H. Schreier, V., & Featherston, S. (2019). Standard items for English judgement studies: Syntax and semantics. In S. Featherston, R. Hörnig, S. von Wietersheim & S. Winkler (Eds.), *Experiments in Focus*. Berlin: de Gruyter.

Haegeman, L. (1988/2009). Parenthetical adverbials: the radical orphanage approach. In S. Benjamin, C. Philippa, F. Werner, & M. Claudia (Eds.), *Dislocated elements in discourse: syntactic, semantic and pragmatic perspectives* (pp. 331–347). Presented at the Conference on Dislocated Elements in Discourse, London, UK: Routledge.

Jackendoff, R. (1977). *X-bar Syntax: A Study of Phrase Structure*. Cambridge: MIT Press.

Kamp, H. & Reyle, U. (1993). From discourse to logic. Dordrecht: Kluwer Academic Publishers.

Kempson, R. (2003). Nonrestrictive relatives and growth of logical form. In G. Garding & M. Tsujimura (Eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics (WCCFL)* (pp. 301-314). Somerville: Cascadilla Press.

McCawley, J. D. (1982). Parentheticals and discontinuous constituent structure. *Linguistic Inquiry, 13*(1)*,* 91-106.

Perzanowski, D. (1980). Appositive relatives do have properties. *Proceedings of the NELSX*, 335–368.

Platzack, C. (2000). A complement-of-N0 account of restrictive and non-restrictive relatives: The case of Swedish. In A. Alexiadou et al. (Eds.), *The Syntax of Relative Clauses* (pp. 265-308). Amsterdam: John Benjamins.

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

Roberts, C. (1988). *Modal Subordination and Pronominal Anaphora in Discourse*. Stanford: CSLI.

Ross, J. (1967). *Constraints on Variables in Syntax*. PhD, MIT.

Safir, K. (1986). Relative clauses in a theory of binding and levels. *Linguistic Inquiry, 17*(4)*,* 663-689.

Sells, P. (1985). *Restrictive and Non-restrictive Modification*. Stanford: CSLI.

Smith, C. (1964). Determiners and relative clauses in a generative grammar of English. *Language*, *40*, 37-52.

Smits, R. (1989). Eurogrammar. The Relative and Cleft Constructions of the Germanic and Romance Languages. Dordrecht: Foris.

Stuurman, F. (1983). Appositives and X bar theory. *Linguistic Inquiry, 14*(4)*,* 736-744.

Thompson, S. (1971). The deep structure of relative clauses. In C. Fillmore & D. Langendoen (Eds.), *Studies in Linguistic Semantics* (pp. 78-94). New York: Holt, Rinehart and Winston.

Toribio, A. (1992). Proper government in Spanish subject relativization. *Probus, 4*, 291-304.