

## RESEARCH

## Verbal mismatch in Right-Node Raising

Aoi Shiraishi<sup>1</sup>, Anne Abeillé<sup>1</sup>, Barbara Hemforth<sup>1</sup> and Philip Miller<sup>2</sup><sup>1</sup> Université de Paris, CNRS, LLF, Paris, FR<sup>2</sup> Université de Paris, CLILLAC-ARP, Paris, FRCorresponding author: Anne Abeillé ([abeille@linguist.univ-paris-diderot.fr](mailto:abeille@linguist.univ-paris-diderot.fr))

Right-Node Raising (RNR, Ross 1967; Chaves 2014) has been claimed to require phonological identity between the missing material and the shared element. Our corpus investigations provide examples of RNR with verb form mismatch with and without syncretism in English and French. Two acceptability experiments show that lack of phonological identity does not affect the acceptability of RNR. We argue further that RNR without phonological identity cannot be taken to be a case of cataphoric VP-ellipsis in French and that it should not be analyzed as such in English. As regards the status of RNR with verb form mismatch, two positions are available: either it is considered to be grammatical, in which case the phonological resolution principle of Pullum & Zwicky (1986) does not hold, or it is considered to be ungrammatical but repaired (or “recycled”, Arregui et al. 2006; Frazier 2013). The high acceptability of cases with mismatch compared with ungrammatical controls casts doubt on the applicability of the recycling hypothesis in such cases. In order to account for the broader range of data established by our corpus and experimental results, we propose a new analysis of RNR based on lexeme identity rather than form identity.

**Keywords:** Ellipsis; French; English; Right-Node Raising; Syncretism; Experimental syntax

## 1 Introduction

This paper studies the construction known as “Right-Node Raising” (RNR, also sometimes called “Right peripheral ellipsis”, first described in Ross 1967; see, e.g., Chaves 2014 for numerous references) in which a right peripheral sequence (typically a constituent) is shared by two or more previous (and typically conjoined) phrases as illustrated in the following examples.<sup>1</sup>

- (1) a. John DETESTS ~~spinach~~ and Mary LIKES spinach. (Chaves 2014: 834)  
b. Sandy has ~~been helping us with the job~~ and you have not been helping us with the job. (Pullum & Zwicky 1986: 761)

We will be more specifically interested in cases where there is mismatch between the peripheral sequence and the first item with which it is shared. It is well known that verb form mismatches may occur between the source and the target in other elliptical constructions, such as VP-ellipsis (VPE) (2-a), Gapping (2-b), and Sluicing (2-c).

<sup>1</sup> In examples of RNR in English Chaves (2014) signals the typical contrastive focus on the items sharing the peripheral sequence by using SMALL CAPS. We often strike out the initial (missing) occurrence of the shared material to clarify the intended interpretation and make mismatches (discussed immediately below) apparent.

- (2)
- a. Peter is complaining about the noise, but John won't ~~complain about the noise~~. (Sag 1976: 4, ex (3))
  - b. I want to try to begin to write a novel and Mary ~~wants to try to begin to write~~ a play. (Ross 1970: 250, ex (1c))
  - c. Decorating for the holidays is easy when you know how to ~~decorate~~. (Merchant 2001: 22, ex (30))

RNR has been claimed to be more restrictive, requiring phonological identity between the missing material and the shared element, as is the case in (1), as opposed to (3-a) and (3-b) (the \* judgments are those of the authors cited). Furthermore, Pullum & Zwicky (1986) propose that a syntactic feature conflict can be resolved by the existence of a syncretic form, so that (3-c) is claimed to be grammatical, as opposed to (3-b).

- (3)
- a. \*I LIKE ~~playing guitar~~ and I WILL play guitar. (Chaves 2014: 870)
  - b. \*I certainly will ~~clarify the situation~~, and you already have, clarify/clarified the situation with respect to the budget. (Pullum & Zwicky 1986: 761)
  - c. I certainly will ~~set the record straight~~, and you already have, set the record straight with respect to the budget. (ibid)

Pullum & Zwicky (1986) note that some speakers accept *clarified* (a case of mismatch) in (3-b) and that further empirical work should be pursued to establish the status of such examples, but to the best of our knowledge, this has never been done.

Several analyses have been proposed for RNR. It has been analyzed in terms of movement (Ross 1967), multidominance (McCawley 1988), and deletion under syntactic identity (Kayne 1994). Under their most natural implementations, all of these analyses predict that mismatch should not be possible.

We have conducted parallel corpus studies and acceptability experiments on English and French to investigate whether verb form identity is in fact required by RNR. We concentrate on cases of type (3) where the right peripheral element is a VP complement, which allows us to raise the question of whether it has the appropriate morphosyntactic form for both of the governors that share it, which we call RNR-MATCH, or whether there is a conflict between the governors, so that the right peripheral VP would not be grammatical if it immediately followed the first governor, which we call RNR-MISMATCH.<sup>2</sup> Furthermore, we investigate whether syncretism plays a role in the acceptability of such cases. We compare the two languages, noting specifically that French does not have VPE (Emonds 1977). This will be crucial to establishing the theoretical consequences of our data, since Chaves (2014) has argued that such apparent cases of RNR with mismatch are in fact cases of cataphoric VPE, an analysis which cannot be extended to French and which, we will argue, should not be adopted for English.

<sup>2</sup> Note that RNR-MISMATCH cases where the peripheral element has the form appropriate for the first governor, e.g., (3-b) in the variant with *clarify* after *have*, are hopelessly unacceptable. As pointed out by a reviewer, with a fronted VP the situation is the opposite:

- (i) Clarify/\*Clarified the situation I certainly will, and you already have.

We suggest that this is not actually RNR, but rather VP-fronting in the first conjunct, followed by VPE in the second.

In this paper, we only discuss verb form mismatch, as opposed to preposition and voice mismatch, which have been studied for French in Abeillé et al. (2016); Shiraishi (2018). In the last section, we propose an HPSG analysis, inspired by Chaves (2014), which also accounts for other cases of mismatches (reflexive and polarity mismatches).

Though we will not be making any new proposals on the discourse pragmatics of RNR, it is useful to sum up here what is known on the topic. As discussed by Hartmann (2000); Ha (2007); Chaves (2014), RNR requires a semantic contrast between two items in the sequences (typically conjuncts) sharing the peripheral sequence (e.g., *detests* and *likes* in (1-a) and *has* and *have not* in (1-b)). We call this the “contrast condition on RNR”. In the absence of such contrast, syntactically well-formed RNR are unacceptable, as shown by the clear difference in acceptability between the following examples:

- (4) a. Bill likes ~~the TV show~~, but Mary dislikes the TV show.  
 b. \*Bill likes ~~the TV show~~, and Mary likes the TV show. (Ha 2007: 2, ex (6))

On the other hand, as pointed out by Chaves, following Kentner et al. (2008), this semantic contrast does not always lead to typical contrastive focus intonation.

## 2 Verbal mismatch in English RNR

### 2.1 Corpus study on English

Our corpus study on English shows that RNR with verbal mismatch is attested but does not suggest that there is any clear preference for syncretic forms.

#### 2.1.1 RNR in the Penn Treebank Corpus

Bîlbîie (2013) reports a study of RNR in the Penn Treebank Corpus (PTB, Marcus et al. 1993; Taylor et al. 2003), which combines the Brown corpus (various written), the Switchboard corpus (spoken) and the Wall Street Journal corpus, for a total of approximately 3.4 million words of parsed text. The PTB has the advantage of having syntactic annotations for ellipsis, which made an exhaustive survey of RNR possible. As reported by Bîlbîie, there are 16 occurrences of RNR of the type we are interested in, where the shared element is a VP governed by two *vs*. All of these examples are matched. A closer look at the occurrences shows that the use of the construction is very stereotypical. Specifically, in 10 cases, the right peripheral VP is shared after an “AUX or not AUX” configuration, with identical AUX, as in (5-a); 5 cases involve other combinations of auxiliaries as in (5-b); only one case involves a coordination of (identical) lexical verbs, the contrast being established on the subjects as in (5-c), further discussed in section 4.2 below).

- (5) a. As a result, the OEX insurance may or may not fully protect an investor’s holdings in the event of a market decline. [WSJ]  
 b. The surge brings to nearly 50 the number of country funds that are or soon will be listed in New York or London. [WSJ]  
 c. The Mexican government is trying, and a lot of the larger Mexican businesses are trying to, oh, make themselves Americanized. [SWB]

#### 2.1.2 RNR in broader corpora

The results of this preliminary search might be taken to suggest that RNR with mismatch is unused. But we found occurrences of verbal mismatch without syncretism on the internet, for example:

- (6) Her publicist Max Clifford said: “I think she’s going to be remembered as a young girl who has ~~saved an awful lot of lives~~, and who will, save an awful lot of lives”.  
(news.bbc.co.uk/2/hi/entertainment/)

This led us to explore corpora which are much larger than the PTB, specifically the English Web Corpus 2015 (enTenTen15, 15.7 billion words)<sup>3</sup> and the Corpus of Contemporary American English (COCA, 560 million words).<sup>4</sup> Since they are not annotated for ellipsis, we used various heuristics likely to lead us to relevant occurrences, for instance searching for certain sequences with coordinated auxiliaries requiring conflicting forms for the following verb, e.g., *who has and who will*, *who have and who are*. The details of these sequences and the numbers of occurrences found with and without syncretism in the enTenTen15 are provided in Table 1. The COCA yielded only 3 occurrences, all non syncretic.<sup>5</sup>

In Table 1, only sequences which actually led to relevant occurrences have been included. It appears that coordination with *and* is more frequent than with *or*<sup>6</sup> and

**Table 1:** RNR with mismatch in the enTenTen15 corpus.

Sequences	Number of occurrences with syncretism	Number of occurrences without syncretism
<i>who have/has and who V + to inf</i>	1	10
<i>who have/has or who V + to inf</i>	1	2
<i>who have/has and who will + inf</i>	0	10
<i>who have/has or who will + inf</i>	0	3
<i>who have/has and who are + ing</i>	0	3
<i>who has or who is + ing</i>	0	2
<i>who will or who are + ing</i>	0	2
<i>who should and who are + ing</i>	0	1
<i>who are and who have + psp</i>	0	1
<i>who are or who have + psp</i>	0	3
<i>who is or who has + psp</i>	0	1
<i>who will or who have/has + psp</i>	0	2
Total	2	40

<sup>3</sup> See <https://www.sketchengine.eu/ententen-english-corpus/> and Baroni et al. (2009).

<sup>4</sup> See <http://corpus.byu.edu/coca/> and Davies (2010).

<sup>5</sup> Two reviewers inquired as to whether it was possible to compare the number of occurrences of matched and mismatched RNR in the corpus. Sadly, the large corpora available do not contain sufficient annotations to reliably extract all cases of RNR (or even a reliable majority of them), so that the question cannot be answered at present. However, it was possible to get some idea of the situation by searching for patterns similar to those in Table 1 but leading to matched RNR cases of the same general type, e.g., *who V[AUX] and/or who V[AUX] (not/n't)*. Using these methods, 269 examples of matched RNR were gathered. This number cannot be directly compared to the 42 of Table 1 (since it is not clear that we are gathering comparable sets), but the internal distribution of examples is very telling: out of 269, 247 are cases where there is the same auxiliary in both conjuncts, the second being negated (e.g., *who will or who will not*). Such examples are of course necessarily matched. Their high frequency can be attributed to the fact that they optimally satisfy the contrast condition on RNR. On the other hand numerous matched patterns have no occurrences (e.g., *who will and/or who can*) or very small numbers of occurrences (e.g., *who can and who should*, with a single occurrence), similar to what one finds in Table 1. It thus seems that what drives frequency is not so much the match vs. mismatch status of a configuration, but the degree to which it easily satisfies the discourse conditions on RNR and the degree to which it expresses something plausible and useful.

<sup>6</sup> 26 for *and* vs. 16 for *or*; however the proportion of uses of *or* in these constructions is clearly much higher than is the case for similar cases of coordination overall: searching for *who V and/or who V* in the enTenTen15 gives us 3801 for *and* vs. 870 for *or*.

that certain sequences are more frequent than others for non syntactic reasons (e.g., *have and will* is more frequent than *will and have* because it reflects actual temporal ordering).

Interestingly, we were able to find only two examples of RNR with conflicting verbs exhibiting a syncretic verb form in the enTenTen15, one of which is given in (7). (8) illustrates a typical nonsyncretic case.

- (7) But isn't the real legacy in the children who have ~~come home to loving families~~ and who continue to come home to loving families?
- (8) Parents, who have ~~requested an exemption for their student not to wear the school uniform~~ or who wish to request an exemption for their student not to wear the school uniform, must adhere to the enclosed Dress and Grooming Policy.

The small proportion of syncretic cases may be a consequence of the small number of verbs which have a syncretic form for the base and past participle. There are 30 such verbs in English (*come, become, set, run, put ...*) and they account for 2.43% of the verb tokens in the enTenTen15.

## 2.2 Acceptability experiment on English RNR

In order to investigate to what extent the corpus examples are actually acceptable, we conducted an acceptability judgment experiment.

### 2.2.1 Materials

Our experimental materials were constructed on the basis of two factors, SYNCRETISM and ELLIPSIS. The first is binary, with the values  $\pm$ SYNCRETIC. The second has three values, namely, (i) RNR without verbal mismatch (RNR-MATCH), (ii) RNR with verbal mismatch (RNR-MISMATCH), and (iii) coordination with the VP anaphor *do so* and no ellipsis (NO ELLIPSIS). These are illustrated in (9) and (10) which respectively show syncretic and non syncretic items. We did not vary SYNCRETISM as a within item factor (which would have led to having 6 conditions for each item by crossing the binary  $\pm$ SYNCRETISM factor with the three values of the ELLIPSIS factor) because it was impossible to create a sufficient number of examples where syncretic and non syncretic verbs plausibly entered in the same contexts.

- |      |             |              |  |
|------|-------------|--------------|--|
| (9)  | + SYNCRETIC | RNR-MATCH    | Many people already have started to, and some soon will bet on Catalan independence.                   |
|      | + SYNCRETIC | RNR-MISMATCH | Many people already have, and some soon will bet on Catalan independence.                              |
|      | + SYNCRETIC | NO-ELLIPSIS  | Many people already have bet on Catalan independence, and some will soon do so.                        |
| (10) | -SYNCRETIC  | RNR-MATCH    | Some new hybrid models have started to, and others will continue to appear in the automobile industry. |
|      | -SYNCRETIC  | RNR-MISMATCH | Some new hybrid models already have, and others soon will appear in the automobile industry.           |
|      | -SYNCRETIC  | NO-ELLIPSIS  | Some new hybrid models already have appeared in the automobile industry and others will soon do so.    |



Our materials for mismatch were constructed as simplifications of the naturally occurring corpus examples, whereas those for match were constructed as variants of the mismatch materials.<sup>7</sup>

We included two types of control sentences in the materials, so that we could compare the acceptability of the target items with similar cases that were clearly grammatical (GRAMMATICAL CONTROL) and cases that are ungrammatical only due to the presence of the wrong verb form after an auxiliary (UNGRAMMATICAL CONTROL) i.e., the same type of mismatch as in the RNR cases in the target sentences. Note that in the ungrammatical control cases the experimental sentences were constructed so that the average distance in words between the auxiliary and the incompatible verb form was the same as that of the RNR examples with mismatch. This was done by inserting adverbials between the auxiliary and the following verb, viz., *most certainly never* in the UNGRAMMATICAL CONTROL stimulus below.

- (11) GRAMMATICAL CONTROL        She is someone who will most certainly never opt  
her children out of NY State testing.  
UNGRAMMATICAL CONTROL    \*She is someone who has most certainly never opt  
her children out of NY State testing.

Thus, if the RNR-MISMATCH materials are ungrammatical, we would expect a penalty similar to that exhibited by UNGRAMMATICAL CONTROL materials since the same kind of violation occurs with the same distance between the conflicting forms.

For both the syncretic and the nonsyncretic cases, twelve items of the types exhibited in (9) and (10) respectively were constructed, resulting in a total of 36 (12 \* 3) experimental sentences for each, i.e., 72 in all. For the control cases, twelve items with two conditions as exhibited in (11) were constructed, resulting in a total of 24 control sentences. In addition there were three practice items and 24 fillers from an unrelated experiment.

Experimental items, controls and fillers were distributed into three lists following a Latin square design so that each participant was confronted with all conditions but with each item in only one of the conditions (24 experimental items and 12 control items per participant). The presentation of conditions was balanced across lists.

### 2.2.2 Procedure and participants

The experiment was set up using the Ixet online platform (Drummond n.d.). All participants provided informed consent before starting the experiment. Participants were recruited on Amazon's Mechanical Turk platform (AMT, <https://www.mturk.com/>). For each trial, participants read a sentence on a computer screen and judged its acceptability on a range of 1 (not at all acceptable) to 10 (completely acceptable). Each acceptability judgment was followed by a simple comprehension question to ensure that the subjects were attentive. The duration of the experiment was estimated to be 10 minutes on average. 42 native speakers of English (19 female) participated in the web-based experiment using their own devices and a location of

<sup>7</sup> Our choice of punctuation in the RNR materials is meant to reflect intonational properties. We place a comma after the first conjunct, where there is always a break, but not after the second, where the break is optional. In any case, the choice of punctuation is consistent across the RNR-MATCH and RNR-MISMATCH conditions so that it cannot be a relevant source of differences between them.

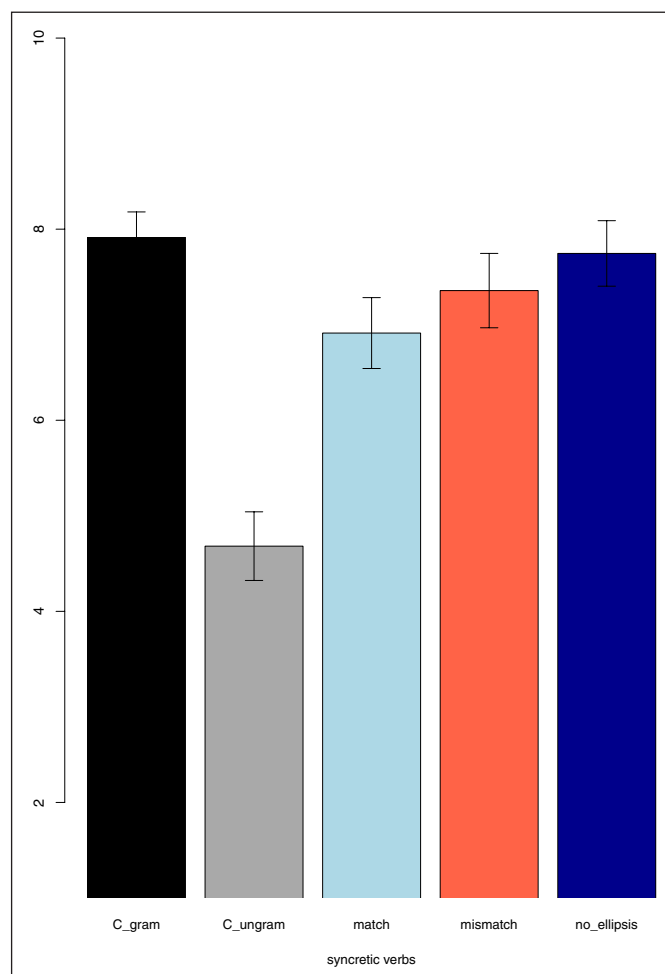
their choice. They were however asked to focus on the experiment in a quiet place and to answer all questions without taking a break. Their age varied between 22 and 61 years old (Mean = 36.7, SD = 9.86). Participants were paid one dollar each for their participation.

### 2.2.3 Results

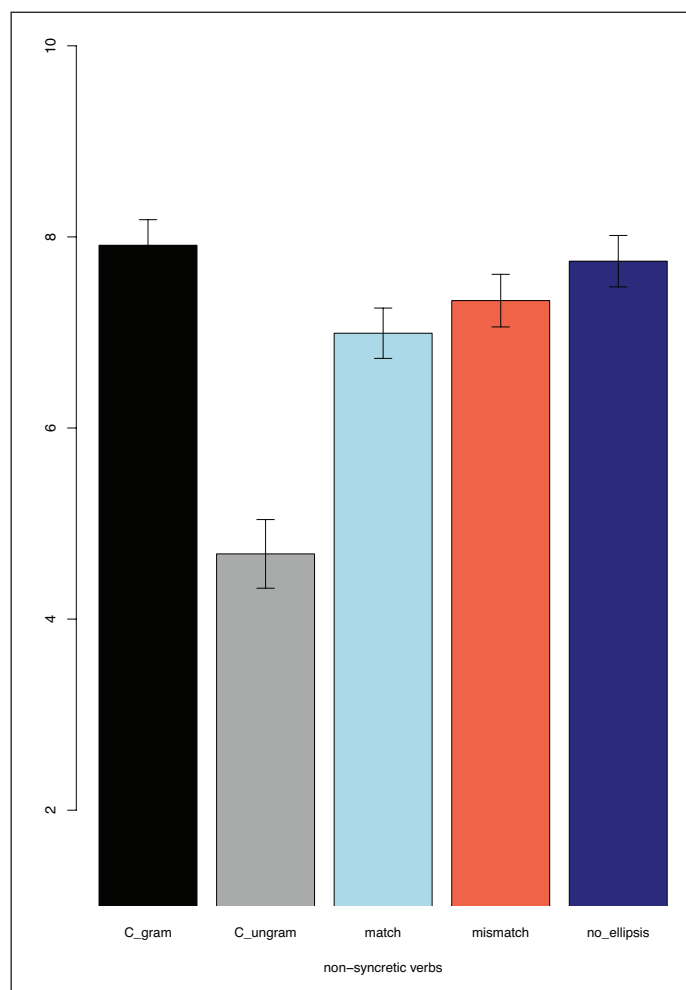
The overall experimental results are shown in Figure 1 (RNR with syncretic forms), Figure 2 (RNR with non syncretic forms). The results for the grammatical and ungrammatical controls have been included in both figures for ease of comparison. Error bars in all figures in this paper correspond to 95 percent confidence intervals.

The details of the statistical analysis can be found in the appendix (Tables 1 to 3). We limit ourselves here to a brief review of what was found to be statistically reliable. Specifically, no main effect nor any statistical interaction was found including the factor *SYNCRETISM*: syncretic forms did not differ in acceptability from non syncretic forms overall, nor did syncretic forms improve the status of RNR with mismatch (as was suggested would be the case by Pullum & Zwicky 1986) as opposed to non syncretic forms (i.e., there was no interaction between *SYNCRETISM* and *RNR-MATCH* vs. *RNR-MISMATCH*).

The *RNR-MISMATCH* and *RNR-MATCH* conditions are significantly less acceptable than the grammatical controls (mismatch:  $p < 0.03$ , match:  $p < 0.001$ ), while the



**Figure 1:** English acceptability experiment: verbs with syncretism and controls.



**Figure 2:** English acceptability experiment: verbs without syncretism and controls.

NO-ELLIPSIS condition (with the VP anaphor *do so* rather than RNR) did not differ significantly from the grammatical controls. On the other hand these three experimental conditions are significantly more acceptable than ungrammatical controls (all  $p$ s < 0.0001). Furthermore, the RNR-MATCH condition is less acceptable than the RNR-MISMATCH condition ( $p$  < 0.05) and than the NO-ELLIPSIS condition ( $p$  < 0.01). The numerical difference in acceptability between the RNR-MISMATCH condition and the NO-ELLIPSIS condition does not reach statistical significance ( $p$  = 0.09459).

#### 2.2.4 Discussion

In our experiment, items with the VP anaphor *do so* instead of ellipsis were judged slightly more acceptable than RNR with or without mismatch. We propose that this is due to the fact that elliptical sentences involve a higher cognitive processing cost than sentences without ellipsis (see, for instance, Roberts et al. 2013 for eye-tracking evidence supporting this idea).

Also, we found no penalty for RNR with mismatch, the match condition was even judged slightly less acceptable. Finally, we found no significant difference depending on whether the verb forms were syncretic or not. These results suggest that syncretic forms do not have a special status in RNR in English and that RNR with verbal mismatch both with and without syncretism should be accounted for similarly by the grammar.



Recall that the conditions with mismatch were constructed as simplifications of the naturally occurring examples found in the corpora, and the match conditions were constructed as variants of the former. This may explain why the match condition was judged slightly less acceptable.<sup>8</sup> However this cannot be the reason for the lack of difference between syncretic and non syncretic mismatch, since both types of examples were inspired from corpus examples (and were not variants of each other).

The question then arises as to whether the acceptability of RNR with mismatch should be interpreted as showing some form of a grammatical illusion (see Frazier 2015). In case of minor errors, the human parser is able to recycle (Grant et al. 2012) or to repair (Frazier & Clifton 2015) degraded input.

We designed our control items to shed some light on this question. Specifically, they exhibit the same verb form mismatches as our experimental target items with mismatch, with the same average distance between the auxiliary trigger and the verb target. They were judged unacceptable, contrary to the RNR cases.

Although Frazier & Clifton (2015) cite cases where what they claim are repaired structures are highly acceptable, their studies of repair in elliptical constructions (e.g., Arregui et al. 2006; Grant et al. 2012) give evidence that repair in ellipsis decreases acceptability.

It might be argued that the acceptability of mismatch constructions is a similar phenomenon to what has been found for attraction errors of the type *the key to the cabinets is/\*are [...]*. Most of the literature on this phenomenon has been concerned with production (see Bock & Miller 1991; Thornton & Macdonald 2003). However certain recent studies (e.g., Reifegerste et al. 2017) use judgment tasks more similar to our experiments (binary grammaticality judgments). They show that attraction to local plural nouns gives rise to grammatical illusions in particular for elderly people and people with low working memory capacity. Similar increases in acceptability of ungrammatical sentences with attraction errors have been found by Schlueter et al. (2018). However, while there was some increase in acceptability for attraction to local plural nouns in both papers, they were still strongly degraded compared to the grammatical variant *the key to the cabinets is*. The acceptability patterns found for attraction errors are thus in no way comparable to those we found for mismatches in our experiments.

Another question to be addressed is whether these cases could be analysed as cataphoric VPE. Chaves (2014) pursues such an approach for apparently acceptable cases of RNR with mismatch. This means that, in an example like (12), the VP in the first conjunct should be analyzed as ellipted by VPE:

- (12) Some new hybrid models already have ~~appeared in the automobile industry~~, and others soon will appear in the automobile industry.

As already mentioned, this idea will not carry over to the similar French data that will be discussed in the next section. But even if one limits one's perspective to English, we argue that, although some cases of apparent RNR may in fact be cases of cataphoric VPE, there are clearly other cases where such a reanalysis lacks plausibility for two reasons.

First, there are intonational differences between the two constructions. Compare (12) and a classical case of cataphoric VPE like (13):

<sup>8</sup> A reviewer suggests that the slight decrease in acceptability of the match cases is linked to the fact that they often involve a stranded infinitival *to*, violating phonological phrasing constraints. Our match materials included 20 items with *to* and 4 without *to*. Although it was impossible to run a full statistical analysis due to the small number of cases without *to*, we found the same numerical pattern: even for items without *to* the mismatch was numerically judged more acceptable than the match condition.

- (13) And she has told me that when I ask her to ~~make it seem to all of the framlings that we here on Lusitania have cut off our ansible connection~~, she can make it seem to all of the framlings that we here on Lusitania have cut off our ansible connection. (O.S. Card, *Speaker for the Dead*, 1986)

As discussed by Kentner et al. (2008), short RNR targets tend to be integrated into the intonational phrase (IP) of the second conjunct whereas long targets will make up their own IP. In (12), where the target is clearly long, there is preferably an intonational break between *will* and the following VP. In (13), on the other hand, it is simply impossible to have a similar intonational break between *can* and the following long VP. It thus seems that RNR and cataphoric VPE have different intonational properties, at least when the target is long.

Beyond this intonational difference, Chaves' suggestion also raises problems with respect to classical constraints on cataphora. It is generally assumed that cataphoric VPE follows the classical constraints on backwards anaphora (Langacker 1969; Reinhart 1983):

- (14) a. Paul<sub>i</sub> left and he<sub>i</sub> was happy.  
 b. #He<sub>i</sub> left and Paul<sub>i</sub> was happy.  
 c. When he<sub>i</sub> left, Paul<sub>i</sub> was happy.

Thus, under these classical analyses, cataphoric VPE should only occur when the anaphor is in a preceding subordinate clause, as in (15-a), as opposed to (15-b). Consequently, coordinations such as (16) could only be analyzed as RNR.

- (15) a. If you can ~~come tomorrow~~, you should come tomorrow.  
 b. #You should ~~come tomorrow~~, if you can come tomorrow.

- (16) If you're scared, you can, and you should, leave now. (Frank Gallagher, John M Del Vecchio, *The Bremer Detail*, 2014)

It has been shown, however, that these strictly syntactic constraints are inadequate. For instance, Mittwoch (1983); Harris & Bates (2002) argue that pragmatic subordination is sufficient to allow cataphoric uses. For instance, Mittwoch's first example is (17):

- (17) I haven't seen *him* yet but **John** is back. (Mittwoch 1983: 131, her markings)

What makes (17) acceptable is the concessive *but*, which makes it pragmatically similar to (18), which satisfies the classical syntactic conditions on cataphora.

- (18) Though I haven't seen him<sub>i</sub>, John<sub>i</sub> is back.

However, as Mittwoch (1983): 130 notes, examples of type (17) are only acceptable if the pronoun is not a first mention of the referent in the discourse. This is corroborated by the study reported in Carden (1982). He collected a large number of cases of cataphora from literary sources, where the cataphor is the first mention of the referent in the novel (as is the case for (13) above). It is remarkable that all of the numerous examples provided by Carden satisfy the classical syntactic constraints. Thus, it appears that these constraints can be relaxed under certain discourse pragmatic conditions, but only in cases where the cataphor is not the first mention of the referent. Thus, an example like (16) might be analyzed as cataphoric VPE if *leave now* was discourse old. However, there is absolutely

no general constraint requiring that the referent of the shared constituent be discourse old in apparently mismatched cases of RNR. E.g., in (12), there is clearly no discourse requirement that *appear in the automobile industry* must have been previously evoked in the discourse. Such cases can only be analyzed as mismatched RNR.

In conclusion, English RNR allows verbal form mismatches and there is no privilege for syncretic forms, contrary to the suggestion made by Pullum & Zwicky (1986).

### 3 Verbal mismatch in French RNR

We now turn to French, which, like English, exhibits Right-Node Raising (Abeillé & Mouret 2010; contra Haspelmath 2007):

- (19) J' ai eu à traiter ~~un certain nombre de dossiers de ce type~~ et je traite  
 I have had to treat ~~a certain number of files of this type~~ and I treat  
 encore un certain nombre de dossiers de ce type.  
 still a certain number of files of this type  
 'I had to deal with and I still deal with a certain number of such files.' (Frank Méjean, *Le guide du divorce*)

In French, in order to relevantly address the question of syncretism in cases of mismatch, one must take care to distinguish spelling and pronunciation. Specifically, verbs of 1st conjugation class have orthographically different endings for the infinitive (-er) and past participle (-é) but these have the same pronunciation (/e/), as illustrated for *fermer* in (20) below. This group covers more than 80% of French verbs, and differs from the other conjugation classes which have non syncretic infinitival and participial forms both in spelling and in pronunciation, as is the case for *voir* in (21) below (specifically, in the other conjugation classes, the final orthographic -r of the infinitive is always pronounced). It is important to note in this context that even highly literate speakers of French regularly make spelling mistakes with 1st conjugation class verbs in this respect, especially in informal contexts (e.g., writing email messages). When discussing French, "syncretism" will always be understood as syncretism in pronunciation. Abeillé & Mouret (2010) mention an attested example of RNR with verbal mismatch from 1st group verb (i.e., with syncretism).

- (20) Certaines agences immobilières ont déjà ~~fermé leurs portes~~, ou vont  
 certain agencies real.estate have already, ~~closed their doors~~, or will  
 bientôt fermer leurs portes.  
 soon close their doors  
 'Some real estate agencies have already, or will soon close their doors.' (*Le Monde* newspaper)

We also found examples without syncretism on the internet.

- (21) ... une carte interactive de tous les sites de production à grande échelle  
 ... a map interactive of all the sites of manufacturing at large scale  
 qui ont ~~vu le jour~~, ou qui vont voir le jour dans les mois qui  
 REL have ~~seen the day~~, or REL will see the day in the months REL  
 viennent en France.  
 come in France  
 '...an interactive map of all the large scale manufacturing sites that have, or that will open in the following months in France.'  
 (franceinter.fr, 2015/02/20)

### 3.1 Corpus study on French

Given these initial data, we carried out a search for RNR with mismatch in a large corpus, namely, the frTenTen12 (more than 10 billion words).<sup>9</sup> Using similar heuristics to those used for English (searching for coordinations of conflicting auxiliaries (*être/avoir* + past participle, *aller/pouvoir* + infinitive) we found 51 examples of verbal RNR with mismatch. These include 65% (33/51) of cases with syncretism (*impacté/impacter* ‘impact’), as in (22-a), and 35% (18/51) without syncretism (*investi/investir* ‘invest’), as in (22-b).

- (22) a. Parler de sujets scientifiques, des innovations qui ont  
 talking of subjects scientific, INDF.PL innovations REL have  
~~impacté le quotidien du grand public~~ ou qui vont impacter  
~~impacted the daily.life of.det.m.sg broad public~~ or REL will impact  
 le quotidien du grand public.  
 the.daily.life of.DEF.M.SG large public  
 ‘Talking about scientific topics, innovations that have or that will impact  
 the daily life of the population at large.’  
 (<http://www.cnrs.fr/centre-est>)
- b. Parmi les nominés, on retrouve les artistes qui ont  
 among the nominees, we find the artists REL have  
~~investi les scènes de France~~ ou qui vont investir les scènes de France.  
~~invested the scenes of France~~ or REL will invest the scenes of France  
 ‘Among the nominees, we find the artists who have or who will invest the  
 French scenes.’  
 ([www.etudiant-france.info](http://www.etudiant-france.info))

The results are provided in Table 2.

As was done above in Table 1 for English, only patterns returning a result were included. As in English, coordinations where textual order corresponds to temporal sequencing are more frequent than the reverse. There are more examples with *ou* (‘or’) than with *et* (‘and’).<sup>10</sup>

**Table 2:** RNR with mismatch in the frTenTen12 corpus.

Sequences	Number of occurrences with syncretism	Number of occurrences without syncretism
<i>qui ont et qui vont + inf</i>	3	2
<i>qui a et qui va + inf</i>	2	1
<i>qui ont et qui peuvent + inf</i>	0	1
<i>qui a et qui peut + inf</i>	1	0
<i>qui ont ou qui vont + inf</i>	19	9
<i>qui a ou qui va + inf</i>	1	0
<i>qui ont déjà ou qui vont + inf</i>	2	1
<i>qui a déjà ou qui va + inf</i>	1	0
<i>qui vont ou qui ont + psp</i>	4	3
<i>qui peuvent ou qui ont + psp</i>	0	1
Total	33	18

<sup>9</sup> See <https://www.sketchengine.eu/frtente-french-corpus/>.

<sup>10</sup> As for English (see fn. 6), we compared to the general case of coordinations of the type *qui v et/ou qui v*, for which we found 30,795 for *et* and 7,241 for *ou*, so that the overrepresentation of *ou* in RNR is even more striking than for English.

Although we found more cases of mismatch with syncretic forms, the non syncretic forms are not marginal, and were found in carefully written texts, such as newspapers or official websites from public institutions. Comparing with our results for English in Table 1, the higher number of cases with syncretic forms can be assumed to follow from lexical distribution: in French more than 90% of verbs have syncretic infinitival and past participle (*-er*, *-é*) whereas in English there are only 30 such verbs.<sup>11</sup>

### 3.2 Acceptability experiment on French RNR

To investigate French RNR in a more controlled environment, we carried out an acceptability judgment experiment, parallel to the one carried out for English.

#### 3.2.1 Materials

Experimental materials follow the same design as for English. Items are either +SYNCRETIC, as in (23), or –SYNCRETIC, as in (24). In each case the factor ELLIPSIS has the three values RNR-MATCH, RNR-MISMATCH, and NO ELLIPSIS (i.e., simple VP coordination with the remnant as object of the first conjunct and a coreferential pronoun as object of the second conjunct). As in English, the RNR-MATCH materials are based on examples from the corpus and the two other conditions are constructed so as to be parallel.<sup>12</sup>

(23)	+ SYNCRETIC	RNR-MATCH	C'est le traitement automatique du langage, qui a pu, et qui va encore bouleverser le courrier électronique en facilitant la traduction.
	+ SYNCRETIC	RNR-MISMATCH	C'est le traitement automatique du langage qui a, et qui va encore bouleverser le courrier électronique en facilitant la traduction.
	+ SYNCRETIC	NO-ELLIPSIS	C'est le traitement automatique des langues qui a bouleversé le courrier électronique, et qui va encore le bouleverser en facilitant la traduction.
(24)	–SYNCRETIC	RNR-MATCH	Quelques électeurs auront bientôt, ou ont peut-être déjà rejoint le centre.
	–SYNCRETIC	RNR-MISMATCH	Quelques électeurs vont bientôt, ou ont peut-être déjà rejoint le centre.
	–SYNCRETIC	NO-ELLIPSIS	Quelques électeurs vont bientôt rejoindre le centre, ou l'ont peut-être déjà rejoint.

As in the English experiment we created control items with two conditions, grammatical controls and ungrammatical controls with an incorrect verb form exhibiting the same type of problem as in the mismatched RNR materials (25). As in English, we added adverbs to make the distance between the auxiliary and the verb similar to the RNR examples. Half of the control sentences involved non syncretic verbs and half involved syncretic verbs (i.e., for the ungrammatical controls, a spelling mistake where *-er* appears instead of *é* or vice versa).

<sup>11</sup> Following a reviewer's suggestion, we also checked the proportion of 1st conjugation class verbs. Since the frTenTen15 is not lemmatized, we used frWaC (Baroni et al. 2009) instead, and found that more than 41% of the verbal tokens belong to the *-er/é* class (Olivier Bonami, pc). This should be compared to the situation for English: in the enTenTen15, as mentioned above, we found that only 2.43% of verbal tokens are from verbs where the base form and past participle are syncretic.

<sup>12</sup> Glosses of the example sentences used in the French materials can be found in Appendix 2.



- (25) GRAMMATICAL CONTROL Les alpinistes n'ont malheureusement pas encore gravi le sommet avant la nuit.  
 UNGRAMMATICAL CONTROL \*Les alpinistes n'ont malheureusement pas encore gravir le sommet avant la nuit.

The number of items of each type was fully parallel to the English experiment: 12 items of type (23), 12 items of type (24), i.e., 72 experimental sentences; 12 items of type (25), i.e., 24 control sentences. In addition there were 36 fillers from an unrelated experiment. Experimental items, controls and fillers were distributed across three lists using a Latin square design so that each list contained 24 experimental sentences, 12 controls, 36 fillers and 3 practice items.

### 3.2.2 Procedure and participants

These materials were included in an acceptability judgment paradigm on the Ixet online platform (Drummond n.d.). After providing informed consent, participants read each sentence on a computer screen and judged its acceptability on a range of 1 (not at all acceptable) to 10 (completely acceptable). Each acceptability judgment was followed by a simple comprehension question to ensure that the subjects were attentive. The duration of the experiment was estimated at 10 minutes on average. 43 native speakers of French (18 to 70 years old, mean = 32.7, SD = 15.07; 23 female), recruited on the RISC website (<http://www.risc.cnrs.fr/>) volunteered to participate in the experiment.

### 3.2.3 Results

The results for syncretic and non syncretic cases are given in Figures 3 and 4. Results for the grammatical and ungrammatical controls have been included in both figures for ease of comparison.

The details of the statistical analysis can be found in the appendix (Tables 4 to 6). Only the central statistically reliable effects are reported here. Similar to the English experiment, no main effect nor any statistical interaction was found including the factor SYNCRETISM: syncretic forms did not differ in acceptability from non syncretic forms overall, nor did syncretic forms improve the status of RNR with mismatch as opposed to non syncretic forms (i.e., there was no interaction between SYNCRETISM and RNR-MATCH vs. RNR-MISMATCH).

All three conditions for the factor ELLIPSIS are slightly but significantly less acceptable than the grammatical controls (all  $ps < .001$ ). On the other hand these three experimental conditions are significantly more acceptable than ungrammatical controls (all  $ps < 0.0001$ ). Furthermore, the RNR-MISMATCH condition is marginally less acceptable than the MATCH condition ( $p = 0.0613$ ) and than the NO-ELLIPSIS condition ( $p < 0.03$ ). The difference in acceptability between the MATCH condition and the NO-ELLIPSIS condition is statistically not reliable ( $p = 0.2464$ ).

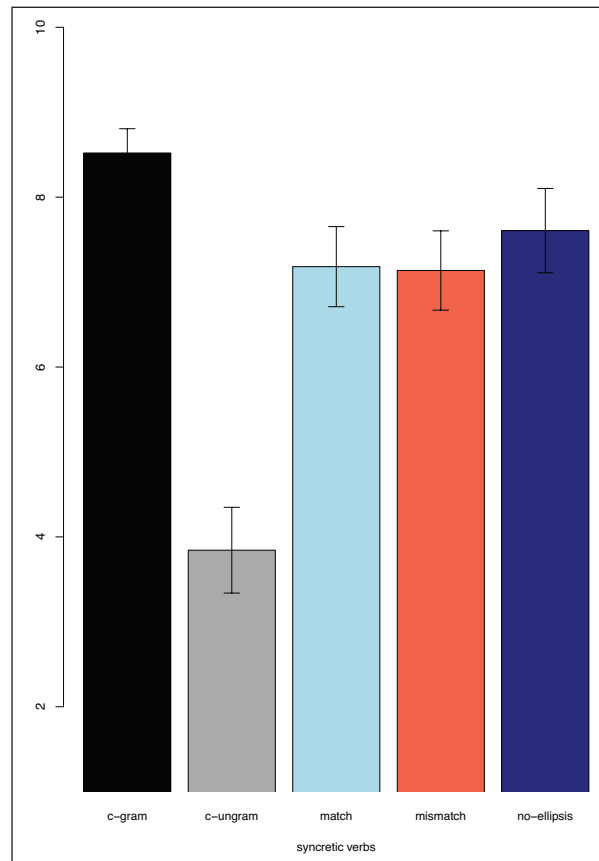
### 3.2.4 Discussion

It is important to note that our RNR examples cannot be analyzed as VPE. Contrary to English, French does not have VPE with tense auxiliaries (Emonds 1977; Abeillé & Godard 2002), be they followed by a past participles (*avoir*, *être*) or by an infinitive (*aller*).<sup>13</sup>

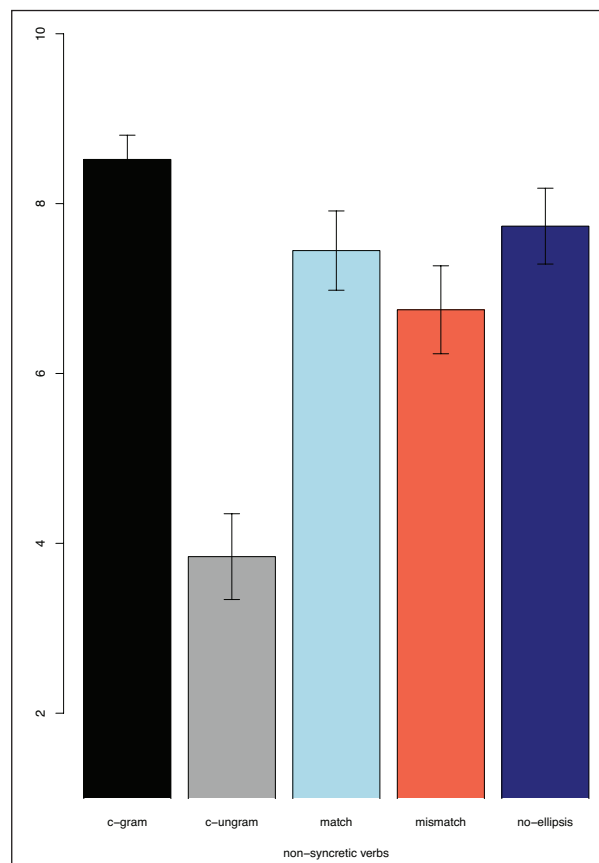
- (26) a. \*Jean n' est pas arrivé, mais Marie est.  
 Jean NEG is NEG arrived, but Marie is  
 'Jean hasn't arrived, but Marie has.'

<sup>13</sup> It has been argued that French allows VPE after some modal verbs (see, e.g., Busquets & Denis 2001; Dagnac 2010) but there is a general consensus that it is impossible after tense auxiliaries, which is why we chose these in constructing our materials.





**Figure 3:** French RNR: verbs with syncretism and controls.



**Figure 4:** French RNR: verbs without syncretism and controls.

- b. \*Jean a fini son travail, mais Marie n' a pas.  
 Jean has finished his work, but Marie NEG has NEG  
 'Jean has finished his work, but Marie hasn't.'
- c. \*Paul va aller à la piscine, mais Marie ne va pas.  
 Paul will go to the pool, but Marie NEG will NEG  
 'Paul will go to the pool, but Marie won't.'

Just as in English, items without ellipsis are slightly more acceptable than RNR with and without mismatch (though only significantly so for mismatch), presumably because elliptical sentences involve higher cognitive processing costs than sentences without ellipsis. There is only a marginal difference between RNR with and without mismatch and no effect at all of syncretism. Notice that, as in English, the mismatch conditions are inspired from the corpus examples, and the items without mismatch are constructed as variants. But this cannot be the reason for the lack of difference between syncretic and non syncretic mismatch, since both types of examples were inspired from corpus examples (and were not variants of one another). The experiment suggests that syncretic forms do not have a special status in RNR in French and that RNR with verbal mismatch with and without syncretism should be integrated in the grammar.

We conclude that both in English and in French RNR should allow for the sharing of different verb forms, like other kinds of ellipsis.

## 4 An HPSG analysis

### 4.1 Previous analyses

There are three major approaches to Right-Node Raising. One of these is rightward movement (Ross 1967; Hankamer 1971; Postal 1974; Gazdar 1981; Steedman 1996; Sabbagh 2007; Sabbagh 2014). This analysis assumes that the right periphery element is outside the coordinate structure and the element is displaced by across-the-board extraction out of the coordinate structure.

(27) [John wrote <sub>-i</sub> and Mary read <sub>-i</sub>] the book<sub>i</sub>.

The second is a multidominance analysis (McCawley 1988; Goodall 1987; Moltmann 1992; Wilder 1999; Gracanin-Yuksekk 2007; Bachrach & Katzir 2009). Such analyses assume that the right periphery element is in the coordinate structure and is dominated by multiple conjuncts.

(28) [John wrote (the book)][and Mary read (the book)].

The third approach is deletion (Wexler & Culicover 1980; Kayne 1994; Hartmann 2000; Yatabe 2001; Beavers & Sag 2004; Yatabe 2012; Chaves 2014). The right periphery element is in the coordinate structure and the right periphery elements in the non-final conjuncts are deleted under certain conditions.

(29) [John wrote ~~the book~~] [and Mary read the book].

Previous analyses of Right-Node Raising in terms of movement, multidominance or deletion do not predict verbal mismatch between the missing material and the shared material.<sup>14</sup>

<sup>14</sup> However, Citko (2018) proposes a multidominance analysis adapted for closest conjunct agreement, which might be extended to the cases of mismatch discussed here.

Even though a movement analysis could be reconciled with mismatch if morphophonology (spell out) applies after movement (Chomsky 1993), various arguments have been proposed against movement analyses (for overall discussion, see Beavers & Sag 2004; Bachrach & Katzir 2009; a.o.), namely RNR can target more than one constituent (Abbott 1976), as shown in (30-a); it can target non constituents, as shown in (31-b) and (31-c) below; and it does not obey locality constraints otherwise observed for wh-movement or extraposition (Levine 1985), as shown in (30-b).

- (30) a. Smith loaned, and his widow later donated, [a valuable collection of manuscripts] [to the library]. (Abbott 1976: 639, ex (1))  
 b. John gave a briefcase, and Harry knows someone who had given a set of steak knives, [to Bill]. (Levine 1985: 492, ex (1))

For all these reasons we do not adopt a movement-based analysis. This means that Right-Node Raising is a misnomer and Peripheral Ellipsis would be a more appropriate name. However we keep the traditional name and acronym here.

#### 4.2 An HPSG analysis of RNR

Constraint based grammars such as Head-driven Phrase Structure Grammar (HPSG, Sag et al. 2003) have proposed a fragment-based analysis (without movement nor deletion) for other kinds of ellipsis such as sluicing (Ginzburg & Sag 2000) and gapping (Abeillé et al. 2014; see Ginzburg & Miller 2019 for an overall perspective on ellipsis in HPSG), thus favoring a direct interpretation approach (see also Culicover & Jackendoff 2005; Culicover & Jackendoff 2012).

However, for RNR, such an approach is difficult due to the high variability of the elliptical material, which does not necessarily correspond to a syntactic constituent (see (30-a), (31-b) and (31-c)). Several versions of a deletion analysis of RNR have been proposed within HPSG, starting with Yatabe (2001). The most recent and most fully developed analysis is that of Chaves (2014), which claims the missing and peripheral strings must have the same morphophonology and that the latter must be prosodically independent. Chaves (following Hudson 1976) has also shown that RNR is not limited to clausal coordination (as illustrated in (31-a), an example from the Brown Corpus mentioned by Bîlbîie 2013) and may involve subconstituents (as in (31-b) and in (30-b) above) and even word parts (as in (31-c)):<sup>15</sup>

- (31) a. She learns how to relax them to accept instead of contracting them to repel the entering object.  
 b. It was a sweet — and an intelligent dog (Switchboard corpus)  
 c. These events took place in pre- or in post-war Germany? (PennTreeBank)

<sup>15</sup> Deletion-based analyses (as well as multidominance) have been argued to be at odds with the possible occurrence in the right periphery of cumulative agreement, as in (i-a), and relational modifiers, as in (i-b) (see Abbott 1976; Jackendoff 1977; Postal 1994).

- (i) a. The pilot claimed that the first nurse and the sailor proved that the second nurse were spies.  
 b. John hummed and Mary sang the same tune.

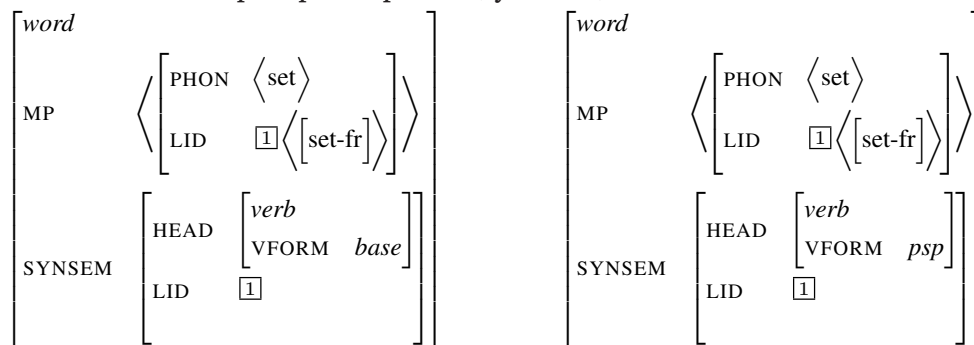
Some have claimed that these cases involve a different kind of elliptical construction (see Yatabe 2003; Valmala 2013 for a tentative distinction between phonological and syntactic RNR). See Chaves (2014) for a more complex version of the deletion analysis which accounts for these facts. Because this aspect of Chaves's analysis is independent from his requirement of morph-form identity, the analysis presented in this paper could easily be extended to include it.

Specifically, Chaves proposes an RNR rule, which targets morphophonological units in Morphophonology (MP). This deletion rule is optional and can target all categories, as well as non-constituents, and can occur in both subordinate and coordinate contexts.

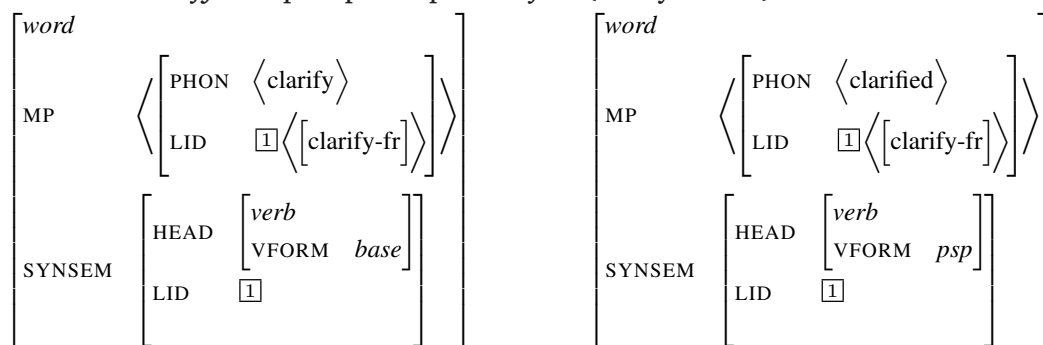
Our analysis is very similar to that proposed by Chaves, but it does not rely on morphophonological identity. Rather, it demands lexeme identity. We use the LID (Lexical Identifier) feature to capture lexeme identity. The LID feature is part of syntactic and semantic features (SYNSEM) and is used to individuate lexical items: it includes a list of semantic frames (represented by “fr” in the feature structures below, where, e.g., set-fr denotes the list of semantic frames appropriate for *set*) that canonically specify the meaning of a lexeme (Sag et al. 2003; Sag 2012). This is used for example for subcategorizing some parts of flexible idioms (such as *pull strings*). Crucially, inflected forms of a lexeme share the same LID feature. Lexeme identity between the infinitive and the past participle is thus captured by the LID feature, which ensures categorical and semantic identity, but not necessarily morphological and phonological identity. Furthermore, we assume that the MP feature includes not only phonological features (PHON) but also a copy of the LID feature.

The feature structures in (32) show simplified lexical entries for the infinitive (BASE) form *set* and the past participle (PSP) form *set*, a case of syncretism. (33) shows simplified lexical entries for the infinitive *clarify* and the past participle *clarified*, where the forms are distinct.

(32) Infinitive *set* and past participle *set* (syncretic):



(33) Infinitive *clarify* and past participle *clarified* (nonsyncretic):



Following Chaves (2014), we assume that RNR is a construction with a mother phrase and a daughter phrase as in (34) (the colors, here and in subsequent figures, have no theoretical significance; they are provided to enhance readability). This rule is optional and may apply to any construction. It is not limited to coordination, but for the sake of simplicity, we will use the terms “non final conjunct” (for the string with ellipsis) and “final conjunct” (for the string with the peripheral element).

$$(34) \left[ \begin{array}{l} mr-cx \\ \text{MOTHER} \left[ \begin{array}{l} \text{MP} \quad [l_1 \oplus r_1 \oplus r_2 \oplus r_3] \\ \text{SYNSEM} \quad [0] \end{array} \right] \\ \text{DTRS} \left\langle \begin{array}{l} \text{MP} \quad [l_1 \oplus l_2 \oplus \langle [\text{PHON} / [p_1]] \dots [\text{PHON} / [p_n]] \rangle] \oplus [r_1 \oplus r_2 \oplus \langle [\text{PHON} / [p_1]] \dots [\text{PHON} / [p_n]] \rangle] \oplus [r_3] \\ \text{SYNSEM} \quad [0] \end{array} \right\rangle \end{array} \right]$$

The MP feature of the Daughter phrase is divided into 5 sublists, which must obey prosodic constraints, which we ignore here. The left hand sublists,  $l_1$  and  $l_2$  (respectively in black and in grey in (34)), correspond to any non final conjunct, and the right hand sublists,  $r_1$  (in blue),  $r_2$  (in red) and  $r_3$  (in black), to the final conjunct. The first sublist  $l_1$  is kept. The deleted sublist  $l_2$  (in grey) must comprise elements with the same LID as  $r_2$  (the shared peripheral element(s), in red). Note that the elements in  $l_2$  are not preserved in the MP feature of the mother phrase (resulting in their not being pronounced). Thus, the form of the peripheral elements  $r_2$  (in red) is always that required by the final conjunct.  $r_1$  (in blue) is the sublist before the shared elements and may comprise a coordinating conjunction. The extra  $r_3$  sublist accounts for Right-node Wrapping and can be empty.<sup>16</sup>

Rule (34) is not limited to coordination and will account for cases such as example (31-a):  $l_1$  is *She learns how to relax them to accept*,  $l_2$  and  $r_2$  are *the entering object*,  $r_1$  is *instead of contracting them to repel*.

Figure 5 provides the analysis for RNR with mismatch and no syncretism, specifically for Pullum and Zwicky’s example (3-b), repeated here:

(35) I certainly will ~~clarify the situation~~ and you already have clarified the situation.

Notice that under our analysis, only LID are shared, so other cases of mismatches are predicted to be acceptable, as long as they share the same lexemes. For example polarity mismatch, in English and French (Abeillé et al. 2016), is accounted for provided *any* and *some* share the same LID in their free choice use, and similarly for French indefinite *un* and *de* ‘a’ (Gaatone 1971; Müller 1991):

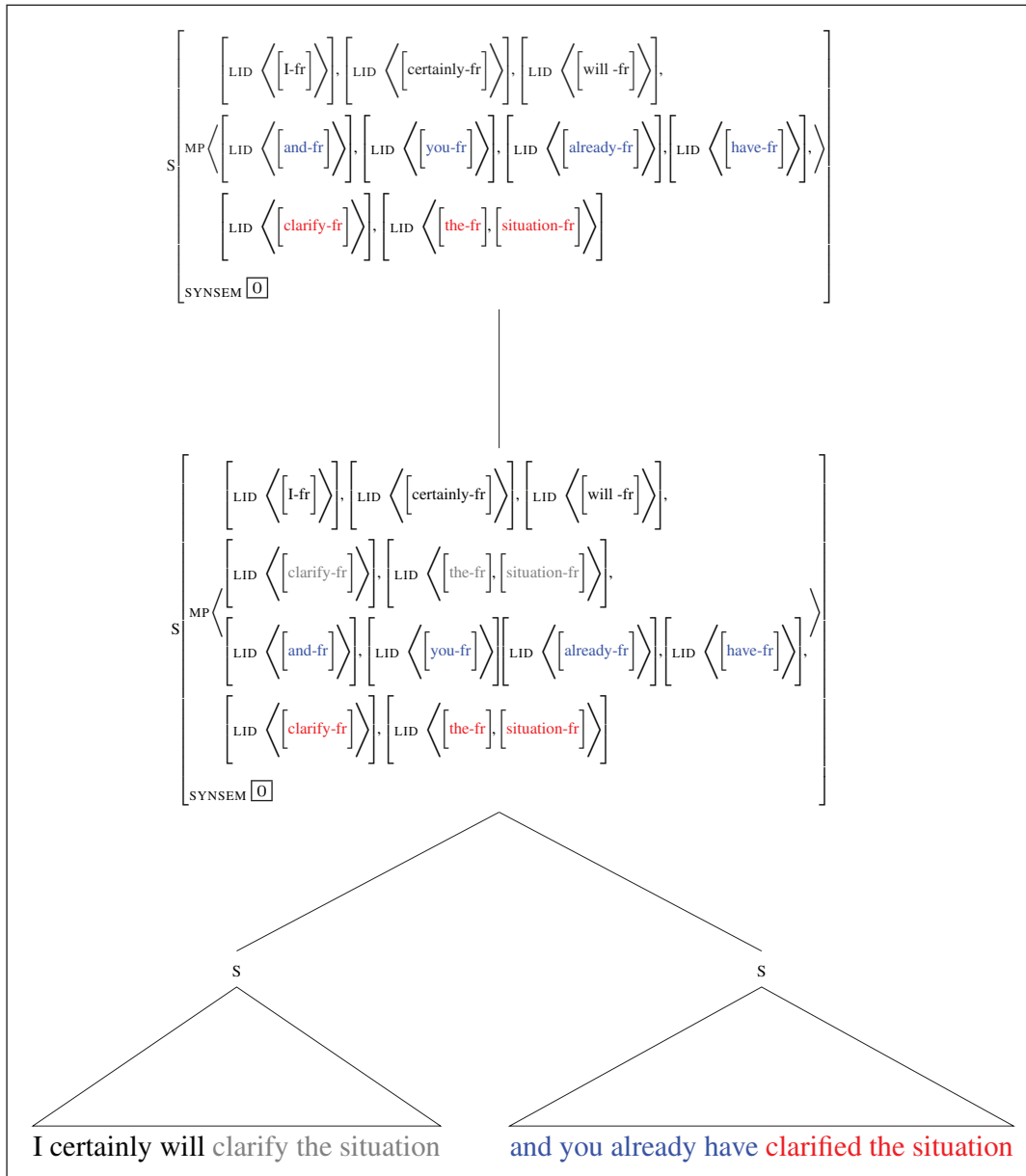
- (36) a. Mary bought ~~some books about linguistics~~, but John didn’t buy any books about linguistics. (Kayne 1994: 146, fn.19)
- b. Il y a des langues qui ont ~~une flexion casuelle~~, et des it there are some languages that have ~~an inflection case-based~~ and some langues qui n’ ont pas, de flexion casuelle. languages that NEG have NEG PREP inflection case-based ‘There are languages that have and languages that don’t have case inflection.’

Similarly, cases of reflexive mismatches are predicted to be grammatical, provided *itself*, *himself*, *themselves* ... share the same LID, and we have indeed found an example in the Penn Treebank (SWB corpus):<sup>17</sup>

<sup>16</sup> See Whitman (2009) who provides numerous attested examples, among which (i), in which the “wrapped” segment *away* corresponds to  $r_3$ :

- (i) I’ve got friends in low places,  
where the whiskey drowns ~~my blues~~ and the beer chases  
*my blues* away.  
 (“Friends in Low Places”, Earl Bud Lee and DeWayne Blackwell, Whitman 2009: 235)

<sup>17</sup> As pointed out by a reviewer, our analysis also predicts closest conjunct agreement (CCA), which has independently been shown to be more acceptable than attraction errors (Keung & Staub 2018). But contrary to CCA, our analysis is not dependent on, nor limited to, coordination.



**Figure 5:** Analysis of English RNR with verbal mismatch without syncretism.

(37) The Mexican government is trying ~~to make itself~~ **Americanized**, and a lot of the larger Mexican businesses are trying to, oh, make themselves Americanized.

Obviously, we are not claiming that any case of mismatch that satisfies the lexical identity constraint will be acceptable.<sup>18</sup> RNR with mismatch is clearly subject to the same complex discourse and prosodic constraints (which go beyond the scope of this paper) that govern RNR with match. For instance, as mentioned above, Chaves’s example given in (3-a), repeated here as (38-a), is probably unacceptable because it doesn’t satisfy the

<sup>18</sup> For example, compare (37) to Chaves 2014’s (42):

(i) \*Tom COULDN’T even though Sue COULD save herself.

We agree that (i) is unacceptable, perhaps because there is a much more natural alternative with VPE: *Tom couldn’t save himself even though Sue could.*



contrast condition on RNR, rather than for reasons of morphosyntactic mismatch, since the matched variant in (38-b) sounds just as bad:

- (38) a. #I like ~~playing guitar~~ and I will play guitar. [our judgment]  
 b. #I like to ~~play guitar~~ and I will play guitar.

Under our analysis the appropriate judgments for these is thus the crosshatch, rather than the asterisk.

## 5 Conclusion

Right-Node Raising has been set apart from other kinds of ellipsis because it was claimed that it forbids mismatch between missing and peripheral elements (with the possible exception of syncretic forms). Through an investigation of English and French corpora, we discovered that examples of RNR with verb form mismatch are attested in both languages, even with nonsyncretic verb forms. Two acceptability studies show that RNR with verb form mismatch is as acceptable as RNR without mismatch, without any preference for syncretic forms, in both languages. Although some cases of RNR with mismatch might be argued to be cataphoric VPE in English, this line of reasoning cannot apply to French, which does not allow VPE after tense auxiliaries. As regards the status of RNR with verb form mismatch, two positions are available: either we consider them to be grammatical, in which case the phonological resolution principle of Pullum & Zwicky (1986) does not hold, or we consider them to be ungrammatical but repaired (or “recycled”). The high acceptability of the cases of RNR with mismatch casts doubt on the applicability of the recycling hypothesis to these cases. Following Chaves (2014), we propose a deletion-based analysis of RNR. In order to account for possible mismatch, we propose to replace morphophonological identity by syntactic and semantic identity defined in terms of lexeme identity.

## Abbreviations

DEF = definite, INDF = indefinite, M = masculine, NEG = negation, PREP = preposition, PSP = past participle, REL = relative, RNR = Right-Node Raising, SG = singular, VPE = Verb-Phrase Ellipsis

## Additional Files

The additional files for this article can be found as follows:

- The scripts and data can be found on: <https://osf.io/x9qmd/>.
- **Appendix.** Statistical analyses for experiment 1 and 2. Translation of French examples (23) and (24). DOI: <https://doi.org/10.5334/gjgl.843.s1>

## Ethics and Consent

Acceptability experiments were run using the Ibex platform. The introductory material before the beginning of the experiments informed the subjects that “Your participation in the experiment is of course strictly voluntary and you are in no way obliged to finish the questionnaire.” In order to actually start the experiment, subjects were furthermore required to check a box in front of the statement “I hereby state that I consent to participate in this experiment.” The following personal information was collected: Age, Gender, highest level of education, native language, region where the native language was learned, other languages spoken. No data allowing the actual identification of individual subjects was collected. Participants received a unique identifier which allows them to retract their consent for usage of their data after the experiment.

## Acknowledgements

We would like to thank the audiences from HeadLex 2016 (Warsaw) and ECBAE 2017 (Lexington) where preliminary versions of this paper were presented, especially Lyn Frazier and Jeff Runner. We would like to thank Rui Chaves and Geoff Pullum for discussion and three anonymous reviewers from *Glossa* for their very useful remarks.

## Funding Information

This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the “Investissements d’Avenir” program (reference: ANR-10-LABX-0083) and contributes to the IdEx Université de Paris (reference: ANR-18-IDEX-0001). Aoi Shiraishi’s PhD research was funded by an “Inspire” grant from the Université de Paris.

## Competing Interests

The authors have no competing interests to declare.

## References

- Abbott, Barbara. 1976. Right Node Raising as a test for constituenthood. *Linguistic Inquiry* 7. 639–642.
- Abeillé, Anne, Berthold Crysmann & Aoi Shiraishi. 2016. Syntactic mismatch in French peripheral ellipsis. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics* 11. 1–30. [http://www.cssp.cnrs.fr/eiss11/eiss11\\_abeille-and-crysmann-and-shiraishi.pdf](http://www.cssp.cnrs.fr/eiss11/eiss11_abeille-and-crysmann-and-shiraishi.pdf).
- Abeillé, Anne & Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78(3). 404–452. <http://muse.jhu.edu/article/19406/pdf>. DOI: <https://doi.org/10.1353/lan.2002.0145>
- Abeillé, Anne & François Mouret. 2010. Quelques contraintes sur les coordinations elliptiques en français. *Revue de sémantique et de pragmatique* 24. 177–207. <http://www.llf.cnrs.fr/Gens/Mouret/abeille-mouret-revised-2011.pdf>.
- Abeillé, Anne, Gabriela Bîlbîie & François Mouret. 2014. A Romance perspective on gapping constructions. In Hans C. Boas & Francisco González-García (eds.), *Romance in Construction Grammar*, 227–267. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/cal.15.07abe>
- Arregui, Ana, Charles Clifton, Lyn Frazier & Keir Moulton. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language* 55. 232–246. DOI: <https://doi.org/10.1016/j.jml.2006.02.005>
- Bachrach, Asaf & Roni Katzir. 2009. Right-Node Raising and delayed spell-out. In Kleantes K. Grohmann (ed.), *InterPhases: Phase-theoretic investigations of linguistic interfaces*, 283–316. Oxford: Oxford University Press. <https://ling.auf.net/lingbuzz/000314/v1.pdf>.
- Baroni, Marco, Silvia Bernardini, Adriano Ferraresi & Eros Zanchetta. 2009. The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation* 43(3). 209–226. DOI: <https://doi.org/10.1007/s10579-009-9081-4>
- Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 48–69. Stanford: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2004/hpsg2004.pdf#page=48>.
- Bîlbîie, Gabriela. 2013. A quantitative study on Right Node Raising in the Penn Treebank. *Presentation at the International Congress of Linguists*, Geneva.

- Bock, Kathryn & Carol A. Miller. 1991. Broken agreement. *Cognitive Psychology* 23. 45–93. DOI: [https://doi.org/10.1016/0010-0285\(91\)90003-7](https://doi.org/10.1016/0010-0285(91)90003-7)
- Busquets, Joan & Pascal Denis. 2001. L'ellipse modale en français: le cas de devoir et pouvoir. *Cahiers de grammaire* 26. 55–74. <https://halshs.archives-ouvertes.fr/halshs-01648372/document>.
- Carden, Guy. 1982. Backwards anaphora in discourse context. *Journal of Linguistics* 18. 361–387. DOI: <https://doi.org/10.1017/S0022226700013657>
- Chaves, Rui Pedro. 2014. On the disunity of Right-Node Raising phenomena: Extraposition, ellipsis, and deletion. *Language* 90(4). 834–886. <https://muse.jhu.edu/article/563570/pdf>. DOI: <https://doi.org/10.1353/lan.2014.0081>
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel J. Keyser (eds.), *The view from building 20*. 1–52. Cambridge, MA: MIT Press.
- Citko, Barbara. 2018. On the relationship between forward and backward gapping. *Syntax* 21(1). 1–36. DOI: <https://doi.org/10.1111/synt.12150>
- Culicover, Peter & Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199271092.001.0001>
- Culicover, Peter W. & Ray S. Jackendoff. 2012. Same-except: A domain general cognitive relation and how language expresses it. *Language* 88(2). 305–340. <https://www.jstor.org/stable/pdf/23251833.pdf?refreqid=excelsior%3Afb67b4077c7d6d2e61f7b368b96c3a5f>. DOI: <https://doi.org/10.1353/lan.2012.0031>
- Dagnac, Anne. 2010. Modal ellipsis in French, Spanish and Italian. Evidence for a TP deletion analysis. In Karlos Arregi, Zsuzsanna Fagyal, Silvina A. Montrul & Annie Tremblay (eds.), *Romance linguistics 2008: Interactions in Romance*, 157–170. Amsterdam: Benjamins. <https://hal-univ-tlse2.archives-ouvertes.fr/hal-00881807/document>. DOI: <https://doi.org/10.1075/cilt.313.15dag>
- Davies, Mark. 2010. The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing* 25(4). 447–464. DOI: <https://doi.org/10.1093/lc/fqq018>
- Drummond, Alex. n.d. Ibex 0.3.8 manual. [http://spellout.net/latest\\_ibex\\_manual.pdf](http://spellout.net/latest_ibex_manual.pdf).
- Emonds, Joseph. 1977. The verbal complex V – V in French. *Linguistic Inquiry* 9(2). 151–175. <https://www.jstor.org/stable/4178050>.
- Frazier, Lyn. 2013. A recycling approach to processing ellipsis. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Diagnosis in syntax*, 485–501. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199602490.003.0024>
- Frazier, Lyn. 2015. Two interpretive systems for natural language? *Journal of Psycholinguistic Research* 44. 7–25. DOI: <https://doi.org/10.1007/s10936-014-9328-0>
- Frazier, Lyn & Charles Clifton, Jr. 2015. Without his shirt off he saved the child from almost drowning: Interpreting an uncertain input. *Language and Cognitive Neuroscience* 30. 635–647. DOI: <https://doi.org/10.1080/23273798.2014.995109>
- Gaetone, David. 1971. *Etude descriptive du système de la négation en français moderne*. Genève: Droz.
- Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12(2). 155–184. <http://www.jstor.org/stable/4178215?origin=JSTOR-pdf>.
- Ginzburg, Jonathan & Ivan Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives*. Stanford, CA: CSLI Publications.
- Ginzburg, Jonathan & Philip Miller. 2019. Ellipsis in Head-Driven Phrase Structure Grammar. In Jeroen van Craenenbroek & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 75–121. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780198712398.013.4>

- Goodall, Grant. 1987. *Parallel structures in syntax: Coordination, causatives, and restructuring*. New York: Cambridge University Press.
- Gracanin-Yukse, Martina. 2007. *About sharing*. Cambridge, MA: MIT dissertation. <https://dspace.mit.edu/bitstream/handle/1721.1/39343/173611094-MIT.pdf?sequence=2>.
- Grant, Margaret, Charles Clifton & Lyn Frazier. 2012. The role of non-actuality implicatures in processing elided constituents. *Journal of Memory and Language* 66. 326–343. DOI: <https://doi.org/10.1016/j.jml.2011.09.003>
- Hankamer, Jorge. 1971. *Deletion in coordinate structures*. New Haven, CT: Yale University dissertation.
- Harris, Catherine L. & Elizabeth A. Bates. 2002. Clausal backgrounding and pronominal reference: A functionalist approach to C-command. *Language and Cognitive Processes* 17. 237–269. DOI: <https://doi.org/10.1080/01690960143000227>
- Hartmann, Katharina. 2000. *Right Node Raising and Gapping: Interface conditions on prosodic deletion*. Amsterdam: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/z.106>
- Ha, Seungwan. 2007. Contrastive focus: Licensor for Right Node Raising. In Emily Elfner & Martin Walkow (eds.), *Proceedings of NELS 37*, 247–260. Urbana-Champaign, IL: University of Illinois. <https://ling.auf.net/lingbuzz/000401/current.pdf>.
- Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Language typology and syntactic description*, 1–50. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511619434.001>
- Hudson, Richard A. 1976. Conjunction reduction, gapping, and right-node raising. *Language* 52. 535–562. <https://www.jstor.org/stable/pdf/412719.pdf>. DOI: <https://doi.org/10.2307/412719>
- Jackendoff, Ray S. 1977. *X-bar syntax. A study of phrase structure*. Cambridge, MA: MIT Press.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: The MIT Press.
- Kentner, Gerrit, Caroline Féry & Kai Alter. 2008. Prosody in speech production and perception: The case of Right Node Raising in English. In Anita Steube (ed.), *The discourse potential of underspecified structures*, 207–226. Berlin: Mouton de Gruyter. [https://www.researchgate.net/profile/Gerrit\\_Kentner/publication/278570480\\_Prosody\\_in\\_Speech\\_Production\\_and\\_Perception\\_The\\_Case\\_of\\_Right\\_Node\\_Raising\\_in\\_English/links/55815bb308aed40dd8cd8e3c/Prosody-in-Speech-Production-and-Perception-The-Case-of-Right-Node-Raising-in-pdf](https://www.researchgate.net/profile/Gerrit_Kentner/publication/278570480_Prosody_in_Speech_Production_and_Perception_The_Case_of_Right_Node_Raising_in_English/links/55815bb308aed40dd8cd8e3c/Prosody-in-Speech-Production-and-Perception-The-Case-of-Right-Node-Raising-in-pdf).
- Keung, Lap-Ching & Adrian Staub. 2018. Variable agreement with coordinate subjects is not a form of agreement attraction. *Journal of Memory and Language* 103. 1–18. DOI: <https://doi.org/10.1016/j.jml.2018.07.001>
- Langacker, Ronald. 1969. On pronominalization and the chain of command. In David A. Reibel & Sanford A. Schane (eds.), *Modern studies in English*, 160–186. New Jersey: Prentice-Hall.
- Levine, Robert. 1985. Right Node (Non-)Raising. *Linguistic Inquiry* 16(3). 492–497. <https://www.jstor.org/stable/4178450>.
- Marcus, Mitchell, Beatrice Santorini & Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn treebank. *Tech. Rep. MSCIS-93-87 University of Pennsylvania Department of Computer and Information Science*. [https://repository.upenn.edu/cgi/viewcontent.cgi?article=1246&context=cis\\_reports](https://repository.upenn.edu/cgi/viewcontent.cgi?article=1246&context=cis_reports). DOI: <https://doi.org/10.21236/ADA273556>
- McCawley, James D. 1988. *The syntactic phenomena of English*. Chicago: Chicago University Press.



- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Mittwoch, Anita. 1983. Backwards anaphora and discourse structure. *Journal of Pragmatics* 7. 129–139. DOI: [https://doi.org/10.1016/0378-2166\(83\)90048-6](https://doi.org/10.1016/0378-2166(83)90048-6)
- Moltmann, Friederike. 1992. *Coordination and comparatives*. Cambridge, MA: MIT dissertation.
- Müller, Claude. 1991. *La négation en français*. Genève: Droz.
- Postal, Paul M. 1974. *On raising*. Cambridge, MA: MIT Press.
- Postal, Paul M. 1994. Parasitic and pseudoparasitic gaps. *Linguistic Inquiry* 25. 63–117.
- Pullum, Geoffrey K. & Arnold M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* 64. 751–773. <https://www.jstor.org/stable/pdf/413900.pdf>. DOI: <https://doi.org/10.2307/415171>
- Reifegerste, Jana, Franziska Hauer & Claudia Felser. 2017. Agreement processing and attraction errors in aging: Evidence from subject-verb agreement in German. *Aging, Neuropsychology and Cognition* 14(6). 627–702. DOI: <https://doi.org/10.1080/13825585.2016.1251550>
- Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. Chicago, IL: The University of Chicago Press.
- Roberts, Leah, Nigel Duffield & Ayumi Matsuo. 2013. Processing VP-ellipsis and VPanaphora with structurally parallel and nonparallel antecedents: An eye-tracking study. *Language and Cognitive Processes* 28. 29–47. DOI: <https://doi.org/10.1080/01690965.2012.676190>
- Ross, John R. 1970. Gapping and the order of constituents. In Manfred Bierwisch & Karl E. Heidolph (eds.), *Progress in linguistics*, 249–259. The Hague: Mouton.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT dissertation.
- Sabbagh, Joseph. 2007. Ordering and linearizing rightward movement. *Natural Language & Linguistic Theory* 25. 349–401. <https://link.springer.com/content/pdf/10.1007%2Fs11049-006-9011-8.pdf>. DOI: <https://doi.org/10.1007/s11049-006-9011-8>
- Sabbagh, Joseph. 2014. Right-Node Raising. *Language and Linguistic Compass* 8(1). 24–35. DOI: <https://doi.org/10.1111/lnc3.12058>
- Sag, Ivan A. 1976. *Deletion and logical form*. Cambridge, MA: MIT dissertation. <https://dspace.mit.edu/bitstream/handle/1721.1/16401/03491413.pdf?sequence=1>.
- Sag, Ivan A. 2012. Sign-based construction grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-based construction grammar*, 69–202. Stanford, CA: CSLI Publications.
- Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. Stanford, CA: CSLI Publications 2nd edn.
- Schlueter, Zoe, Alexander Williams & Ellen Lau. 2018. Exploring the abstractness of number retrieval cues in the computation of subject-verb agreement in comprehension. *Journal of Memory and Language* 99. 74–89. DOI: <https://doi.org/10.1016/j.jml.2017.10.002>
- Shiraishi, Aoi. 2018. *Discordances dans l'ellipse périphérique en français*. Paris: Université Paris Diderot dissertation.
- Steedman, Mark. 1996. *Surface structure and interpretation*. Cambridge, MA: The MIT Press.
- Taylor, Ann, Mitchell Marcus & Beatrice Santorini. 2003. The Penn treebank: An overview. In Anne Abeillé (ed.), *Treebanks*, 5–52. Dordrecht: Springer. DOI: [https://doi.org/10.1007/978-94-010-0201-1\\_1](https://doi.org/10.1007/978-94-010-0201-1_1)

- Thornton, Robert & Maryellen C. MacDonald. 2003. Plausibility and grammatical agreement. *Journal of Memory and Language* 48. 740–759. DOI: [https://doi.org/10.1016/S0749-596X\(03\)00003-2](https://doi.org/10.1016/S0749-596X(03)00003-2)
- Valmala, Vidal. 2013. On Right Node Raising in Catalan and Spanish. *Catalan Journal of Linguistics* 12. 219–251. <https://www.raco.cat/index.php/CatalanJournal/article/download/292433/380967>. DOI: <https://doi.org/10.5565/rev/catjl.64>
- Wexler, Kenneth & Peter W. Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.
- Whitman, Philip Neal. 2009. Right-node wrapping. In Erhard Hinrichs & John Nerbonne (eds.), *Theory and evidence in semantics*, 235–56. Stanford, CA: CSLI Publications.
- Wilder, Chris. 1999. Right-Node Raising and the LCA. In Sonya Bird, Andrew Carnie, Jason D. Haugen & Peter Norquest (eds.), *Proceedings of the 18th West Coast Conference on Formal Linguistics*, 586–598. <http://home.uni-leipzig.de/muellerg/wilder99.pdf>.
- Yatabe, Shûichi. 2001. The syntax and semantics of Left-Node Raising in Japanese. In Dan Flickinger & Andreas Kathol (eds.), *Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, 325–344. Stanford: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2000/hpsg00yatabe.pdf>.
- Yatabe, Shûichi. 2003. A linearization-based theory of summative agreement in peripheral-node raising constructions. In Jong-Bok Kim & Stephen Wechsler (eds.), *The Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 391–411. Stanford: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/3/yatabe.pdf>.
- Yatabe, Shûichi. 2012. Comparison of the ellipsis-based theory of non-constituent coordination with its alternatives. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University Daejeon*, 453–473. <https://pdfs.semanticscholar.org/b047/5b8d1f434ac222a9d18e2f aa76adab20d8d4.pdf#page=454>.

**How to cite this article:** Shiraishi, Aoi, Anne Abeillé, Barbara Hemforth and Philip Miller. 2019. Verbal mismatch in Right-Node Raising. *Glossa: a journal of general linguistics* 4(1): 114.1–26. DOI: <https://doi.org/10.5334/gjgl.843>

**Submitted:** 19 October 2018    **Accepted:** 16 May 2019    **Published:** 18 October 2019

**Copyright:** © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS