

Appendix to Verbal mismatch in Right-Node Raising

Aoi Shiraishi
Université de Paris
CNRS, LLF
5, rue Thomas Mann, F-75013 Paris
ioaaoi430@gmail.com

Anne Abeillé
Université de Paris
CNRS, LLF
5, rue Thomas Mann, F-75013 Paris
abeille@linguist.univ-paris-diderot.fr

Philip Miller
Université de Paris
CLILLAC-ARP
5, rue Thomas Mann, F-75013 Paris
philip.miller@univ-paris-diderot.fr

Barbara Hemforth
Université de Paris
CNRS, LLF
5, rue Thomas Mann, F-75013 Paris
barbara.hemforth@linguist.univ-paris-diderot.fr

1 Statistical analysis for English experiment

The data were analyzed with R version 3.5.0. Ratings were submitted to a linear mixed effects regression (Baayen 2008) using the `lmer`-function from the R-package `lme4`, version 1.1-17 (Bates et al. 2015). We first ran a model to compare RNR variants and verb forms. We used the factor `ELLIPSIS` type (RNR with verbal mismatch, VP anaphor *do so* without ellipsis, RNR without mismatch) as a within participant and within items predictor and the factor `SYNCRETISM` (syncretic, non-syncretic) as a within participant but between items predictor. In addition to random intercepts of items and participants, we included random slopes for `ELLIPSIS` type for items and for participants. Random slopes for `SYNCRETISM` were also included for participants. P-values were computed with the `lmerTest`-package (Kuznetsova et al. 2017), using Satterthwaite approximation of degrees of freedom. Predictors were centered for this model. The VP anaphor *do so* without ellipsis is the reference condition for `ELLIPSIS`. Results are presented in Table 1. Following this general model, we compared the experimental conditions to grammatical (Table 2) and ungrammatical (Table 3) controls using treatment coding, meaning that each of the ellipses conditions are compared to the grammatical or ungrammatical control. In these latter models, random slopes for Condition were included for participants but not for items since grammatical/ungrammatical sentences were constructed as separate items.

Table 1 shows an effect of the match and the mismatch conditions which are slightly less acceptable than the VP anaphor *do so* without ellipsis condition across verb forms. No main effect or interaction was found for `SYNCRETISM`. A separate model, using the *match* condition as the reference condition for `ELLIPSIS` shows that the *mismatch* condition is slightly but significantly more acceptable than the *match* condition (Estimate: 0.35263, Std. Error: 0.16804, df: 45.68633, t:2.098, p=0.04142).

Table 1: Syncretism and Ellipsis type

	Estimate	Std.Error	df	t value	Pr(= t)
General mean for syncretism, no ellipsis	7.35538	0.23217	50.14390	31.681	< 2e-16 ***
syncretism	-0.04649	0.21088	21.77282	-0.220	0.82758
mismatch	-0.40828	0.23560	26.83718	-1.733	0.09459 .
match	-0.76122	0.26409	22.44132	-3.973	0.00854 **
syncretism:mismatch	0.09851	0.44818	23.87181	-0.193	0.84845
syncretism:match	-0.10351	0.53596	24.42637	-0.344	0.73456

Since no main effect or interaction was found including SYNCRETISM, comparisons between experimental items and grammatical and ungrammatical controls were done across syncretic and non-syncretic verbs. Tables 2 and 3 show that match and mismatch RNRs are slightly less acceptable than grammatical controls while sentences with the VP anaphor *do so* but without ellipsis did not significantly differ from the grammatical controls. All three experimental conditions are significantly more acceptable than ungrammatical controls.

Table 2: Comparison to grammatical controls

	Estimate	Std.Error	df	t value	Pr(= t)
grammatical control	7.9127	0.2612	57.8251	30.289	< 2e-16 ***
match	-0.9207	0.2500	47.9003	-3.682	0.000586 ***
mismatch	-0.5793	0.2427	45.5484	-2.387	0.021221 *
no ellipsis	-0.1652	0.2325	47.6163	-0.710	0.480874

Table 3: Comparison to ungrammatical controls

	Estimate	Std.Error	df	t value	Pr(= t)
ungrammatical control	4.6825	0.3515	58.1483	13.323	< 2e-16 ***
match	2.3091	0.4012	58.6934	5.756	3.33e-07 ***
mismatch	2.6512	0.4265	58.3766	6.216	5.86e-08 ***
no ellipsis	3.0656	0.4088	59.2380	7.498	3.79e-10 ***

2 Glosses of the examples of the French materials

- (22) a. C'est le traitement automatique du langage, qui a pu, et
 It's the processing automatic of language, which has could, and
 qui va encore bouleverser le courrier électronique en facilitant
 which will further upset the mail electronic by facilitating
 la traduction.
 the translation.
 'It is automatic language processing, which has been able to, and will
 further revolutionize e-mail by making translation easier.'
- b. C'est le traitement automatique du langage qui a, et qui
 It's the processing automatic of language, which has, and which

- va encore bouleverser le courrier électronique en facilitant la
will further upset the mail electronic by facilitating the
traduction.
translation.
'It is automatic language processing, which has, and which will further
revolutionize e-mail by making translation easier.'
- c. C'est le traitement automatique des langues qui a bouleversé
It's the processing automatic of language, which has upset
le courrier électronique, et qui va encore le bouleverser en
the mail electronic, and which will further it upset by
facilitant la traduction.
facilitating the translation.
'It is automatic language processing, which has revolutionized e-mail,
and which will further upset it by making translation easier.'
- (23) a. Quelques électeurs auront bientôt, ou ont peut-être déjà rejoint
Some voters will have soon, or have perhaps already joined
le centre.
the center.
'Some voters will soon have, or may have already joined the center.'
- b. Quelques électeurs vont bientôt, ou ont peut-être déjà rejoint le
Some voters will soon, or have perhaps already joined the
centre.
center.
'Some voters will soon, or may have already joined the center.'
- c. Quelques électeurs vont bientôt rejoindre le centre, ou l'ont peut-être
Some voters will soon join the center, or it have perhaps
déjà rejoint.
already joined.
'Some voters will soon join the center, or may have already joined it'
- (24) a. Les alpinistes n'ont malheureusement pas encore gravi le
The mountaineers NEG have unfortunately NEG yet climbed the
sommet avant la nuit.
summit before the night.
'The mountaineers have unfortunately not climbed the summit before
dark.'
- b. *Les alpinistes n'ont malheureusement pas encore gravir le
The mountaineers NEG have unfortunately NEG yet climb the
sommet avant la nuit.
summit before the night.

3 Statistical analysis for French experiment

Data analysis was done using the same statistical models as before. Just as for the experiment on RNR in English, we first ran a model to compare RNR variants and

verb forms. ELLIPSIS type (RNR with verbal mismatch, without ellipsis, RNR without mismatch) as a within participant and within items predictor and SYNCRETISM (\pm syncretic) as a within participant but between items predictor. In addition to random intercepts of items and participants, we included random slopes for ELLIPSIS type for items and for participants. Random slopes for SYNCRETISM were also included for participants. P-values were computed with the lmerTest-package (Kuznetsova et al. 2017), using Satterthwaite approximation of degrees of freedom. Predictors were centered for this model. The NO ELLIPSIS condition is the reference condition for ELLIPSIS. To avoid non-convergence problems, we used the optimx optimizer and we excluded correlations of random intercepts and slopes (Nash & Varadhan 2011; Nash 2014; Bates et al. 2015).

Results are presented in (Table 4). Following this general model, we compared the experimental conditions to grammatical (Table 5) and ungrammatical (Table 6) controls using treatment coding, meaning that each of the ellipses conditions is compared to the grammatical or ungrammatical control. In these latter models, random slopes for Condition were included for participants but not for items since grammatical/ungrammatical sentences were constructed as separate items.

Table 4 shows no significant effects of the match condition. The RNR-MISMATCH condition is however significantly less acceptable than NO ELLIPSIS. A model using RNR-MATCH as reference category shows that the RNR-MISMATCH condition is marginally less acceptable (Est.=-0.38770, Std.Error=0.19347, df=16.95952, p=.0613). As for English, no significant main effect or interaction was found for SYNCRETISM.

Table 4: Syncretism and Ellipsis type.

	Estimate	Std.Error	df	t value	Pr(= t)
(Intercept/General mean)	7.300089	0.322840	46.830579	22.612	< 2e-16 ***
syncretism	-0.006964	0.379385	24.534625	-0.018	0.9855
mismatch	-0.822061	0.344024	22.314506	-2.390	0.0257 *
match	-0.425910	0.357736	22.249025	-1.191	0.2464
syncretism:mismatch	0.321799	0.662566	20.840145	0.486	0.6323
syncretism:match	-0.103954	0.685181	20.237625	-0.152	0.8809

Comparisons of experimental items to grammatical and ungrammatical controls were done across syncretic and non-syncretic verbs because no main effect or interaction was found including SYNCRETISM. Tables 2 and 3 show that match and mismatch RNRs are slightly less acceptable than grammatical controls while sentences without ellipsis did not significantly differ from the grammatical controls. All three experimental conditions are significantly more acceptable than ungrammatical controls.

Table 5: Comparison to grammatical controls

	Estimate	Std.Error	df	t value	Pr(= t)
grammatical control	9.3308	0.2797	58.1261	33.359	< 2e-16 ***
match	-2.5031	0.3676	55.6231	-6.809	7.36e-09 ***
mismatch	-2.5163	0.3655	57.7606	-6.885	4.66e-09 ***
no ellipsis	-1.7039	0.3083	65.2294	-5.527	6.13e-07 ***

Table 6: Comparison to ungrammatical controls

	Estimate	Std.Error	df	t value	Pr(= t)
ungrammatical control	3.2821	0.4326	53.8822	7.586	4.66e-10 ***
match	3.5361	0.5377	55.4885	6.576	1.80e-08 ***
mismatch	3.5396	0.5011	52.1992	7.064	3.84e-09 ***
no ellipsis	4.3523	0.5386	54.4006	8.081	6.94e-11 ***

References

- Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1–48.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82. 1–26.
- Nash, John C. 2014. On best practice optimization methods in R. *Journal of Statistical Software* 60. 1–14. <https://www.jstatsoft.org/index.php/jss/article/view/v060i02/v60i02.pdf>.
- Nash, John C. & Ravi Varadhan. 2011. Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software* 43. 1–14. <https://www.jstatsoft.org/index.php/jss/article/view/v043i09/v43i09.pdf>.