Independent Contributions of Semantic Factors to Subject-Verb Agreement Errors during Production

Darrell J. Penta & Neal J. Pearlmutter (Northeastern University)

INTRODUCTION

WHAT TYPES OF INFORMATION AFFECT SUBJECT-VERB AGREEMENT DURING SENTENCE PRODUCTION?

Subject-verb agreement: English conventionally requires the subject and verb of a sentence to be matched on the feature of grammatical number (i.e., singular vs. plural; S vs. P).

Agreement errors: Cases in which the numbers of a subject and verb do not match.

■ INFORMATION INVOLVING NUMBER

Grammatical number (Bock & Miller, 1991)
- Greater error rate in participants' sentence completions when preambles contain a singular head noun (HN) followed by a plural vs. a singular local noun.

Error rate = agreement errors / (agreement errors + correct, number-marked verbs)

Mismatch effect = SP error rate - SS error rate

Noun-types in example preamble			Verb-agreement type in example completion		
	HN	LN	Correct singular	Erroneous plural	Match condition
The	*KEY* *to the*	*CABINET*	*is old*	**are old*	SS
The	*KEY* *to the*	*CABINETS*	*is old*	**are old*	SP

Number meaning (Eberhard, 1999)
- Conceptual number valuations in a speaker's to-be-uttered message modulate the size of the mismatch effect.

■ SEMANTIC INFORMATION NOT INVOLVING NUMBER

Semantic integration (integration; Solomon & Pearlmutter, 2015)
- Reflects how tightly linked the words of a sentence are in a speaker's message.
- Strength of link depends on how the words are construed in the context of a given preamble.
- Larger mismatch effect for integrated vs. unintegrated preambles:

The pizza with the yummy topping(s) (Integrated)
The pizza with the tasty beverage(s) (Unintegrated)

Semantic relatedness (relatedness; Barker, Nicol, & Garrett, 2001)
- Reflects the strength of a connection between two words based on their meaning.
- Strength of relationship is independent of context.
- Larger mismatch effect for related vs. unrelated preambles:

The sailboat by the canoe(s) (Related)
The sailboat by the cabin(s) (Unrelated)

Association
- Reflects the probability of two words cooccurring (see Nelson, McEvoy, & Schreiber; 2004).
- Covaries with relatedness in many cases.

skeleton—bone(s) (Associated/related)
skeleton—closet(s) (Associated)
skeleton—classroom(s) (Unassociated)

HOW DO THESE SEMANTIC FACTORS RELATE TO EACH OTHER ?

Penta & Pearlmutter (2015) Exp.1 (P&P Exp.1) crossed integration and relatedness; Exp. 2 (P&P Exp. 2) varied association with relatedness.

Exp. 1 mismatch effects:
Related > Unrelated
Integrated = Unintegrated

Exp. 2 mismatch effects:
Associated/related > Unassociated
Associated > Unassociated
Related = Unrelated

RESEARCH QUESTIONS

TO WHAT EXTENT DO INTEGRATION, RELATEDNESS, AND ASSOCIATION INDEPENDENTLY CONTRIBUTE TO THE AGREEMENT ERRORS IN PENTA AND PEARLMUTTER (2015)?

■ DOES INTEGRATION INCREASE THE PROBABILITY OF AGREEMENT ERRORS WHEN THE EFFECT OF RELATEDNESS IS HELD CONSTANT?

- The mismatch effect difference was not significant in P&P Exp.1, but the pattern of error rates matched the pattern of normed integration ratings.
- The manipulation was weak due to controlling for a confound with relatedness.

■ HOW DO GRADED DIFFERENCES IN ASSOCIATION AND RELATEDNESS AFFECT AGREEMENT ERRORS?

- ANOVAs of P&P Exp.2 error data revealed no effect of relatedness in the absence of association.
  - A smaller mismatch effect for related associates versus unrelated associates was marginally significant.
- Analyses of the association and relatedness using continuous measures should clarify the nature of the relationship between these factors.

ANALYSES & RESULTS

ERROR DATA FROM PENTA & PEARLMUTTER (2015) WERE ANALYZED USING EMPIRICAL LOGIT-WEIGHTED LINEAR MIXED-EFFECT REGRESSIONS

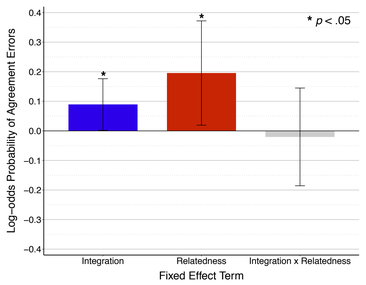
- ANOVAs on both raw and arcsine-transformed proportions for binary outcome variables (e.g., agreement errors) may produce spurious results (Jaeger, 2008)
- Empirical logit-weighted linear mixed-effect regressions are more sensitive to variability in response outcomes as a function of one or more predictors,

- Analyses on log-transformed plural-LN error proportions.
- All models included preamble word-length & frequency control measures.
- All models included random slopes & random intercepts
- All continuous predictors were centered & scaled.

Fixed effects included:
- Exp.1: relatedness x integration, as effects-coded factors.
- Exp.2: continuous relatedness x association measures.

Random effects included:
- Exp.1: main effects terms for relatedness & integration.
- Exp.2: the highest-order interaction term for relatedness & association.

■ BOTH INTEGRATION AND RELATEDNESS INCREASE THE PROBABILITY OF ERRORS WHEN THE EFFECT OF THE OTHER FACTOR IS HELD CONSTANT



■ ASSOCIATION INCREASES THE PROBABILITY OF ERRORS; RELATEDNESS DECREASES THE ODDS; THE EFFECT OF RELATEDNESS DECREASES WITH INCREASED ASSOCIATION

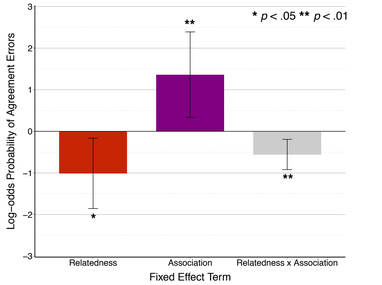


Figure 2. P&P Exp.2 regression results. Relatedness is the scaled and centered mean rating for HN-LN pairings. Association is the scaled and centered arcsine-transformed mean proportion representing the number of times either the HN or LN was elicited during norming for a LN or HN prompt, respectively, out of the total number of responses given for both prompts. Both means were aggregated separately within in combination of item, condition, and experiment list. Error bars are 95% confidence intervals.

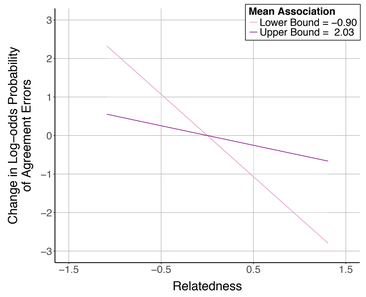


Figure 3. The effect of relatedness conditioned on association for Exp. 2 data, with the factors specified as described in the caption to Figure 2.

DISCUSSION

- ANOVA on these data yielded an effect of relatedness, but no effect of integration. Regression analyses confirmed the relatedness effect but provide evidence of an independent contribution from integration.
- This finding is consistent with numerous published studies showing higher rated integration increasing error rates.
- The regressions also clarified the pattern of results concerning association and relatedness:
  - First, it confirms ANOVAs on these data showing increased error rates as a function of increased relatedness.
  - Second, it provides evidence that increased relatedness decreases the probability of errors when controlling for association. This result warrants further attention, especially as it appears to contradict the findings of Barker et al. (2001), whose manipulation of relatedness purportedly did not involve association.
  - Third, it indicates that the effect of relatedness is attenuated as association increases.
- In as much as association had not been directly tested prior to P&P Exp.2, these result offer additional critical support for the role of this variable on increasing the incidence of agreement errors during production.
- These results do not provide direct evidence about potential mechanisms through which association affects agreement. We are currently investigation the possibility that association can increase error rates through priming.
- It is also unclear what the component dimension of relatedness are separate from association. Manipulations of synonymy— a specific type of relatedness— show clear effect on other language production outcomes. Preliminary analyses of normed ratings of synonymy for the Exps. 1&2 stimuli suggest that synonymy may be a better predictor of decreased error probabilities than relatedness.

REFERENCES

Bock & Miller (1991). *Broken agreement.* Cognitive Psychology, 23, 45–93.
Eberhard (1999). The accessibility of conceptual number to the processes of subject-verb agreement in English. *Journal of Memory and Language*, 41, 560–578.
Penta & Pearlmutter (2015). *Effects of semantic & associative relations on subject-verb agreement.* Manuscript in preparation.
Solomon & Pearlmutter (2004). Semantic integration & syntactic planning in language production *Cognitive Psychology*, 49, 1–46.
Barker, Nicol, & Garrett (2001). Semantic factors in the prod. of number agreement. *Journal of Psycholinguistic Research*, 30, 91–114.
Jaeger (2008). Categorical data analysis: Away from ANOVAs (transformation or not) & towards logit mixed models. *Journal of Memory and Language*, 59, 434–446.
Barr (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*,, 59, 457–474.

ACKNOWLEDGEMENTS

We thank Catherine Balboni, Camila Demaestri, Amy DiBattista, Shreya Divatia, Monika Izdebski, Deanna Marshall, Ian More, Talis Reks, Kesley Thompson, Maria Van Thienen, and J.P. Vangsgaard for their help in collecting and transcribing data.