

Predicting (mis)matches in sluicing: Evidence from cloze, rating and reading time data

Robin Lemke, Lisa Schäfer, Ingo Reich (Saarland University)

The (un)acceptability of mismatches between antecedent and target is often interpreted as evidence for syntactic identity conditions (Merchant, 2013) on ellipsis. We explore whether some of the relevant data can be explained by predictive processing mechanisms (Hale, 2001; Levy, 2008) without assuming specific identity conditions. We investigate this at the case of sluicing: Chung (2006) observes that a mismatch between a PP antecedent and a DP sluice is possible under sluicing (1a), but not under sprouting (1b). Chung accounts for this by proposing an identity condition that requires all words that are omitted in the target (*John, danced, with*) to be given in the antecedent.

- (1) a. John danced with somebody, but I don't know (with) who(m). Sluicing
b. John danced, but I don't know *(with) who(m). Sprouting

Account We hypothesize that such mismatches are not ungrammatical, but that they are degraded because they are unlikely. Unpredictable words are harder to process (Hale, 2001; Levy, 2008) and excessively high processing effort reduces acceptability (Sag et al., 2007). If the overt antecedent in (1a) makes a second participant, and consequently a sluice referring to him/her, more likely than in (1b), this difference in likelihood might result in reduced acceptability. Additionally, our processing account predicts that any manipulation of the likelihood of a sluice will affect its acceptability. For instance, the verb *to dance* might increase the likelihood of a second participant to whom a sluice could refer as compared to e.g. *to present* in *John presented a paper, but I don't know* This is not predicted by a grammatical account like Chung (2006).

Methods We investigated these predictions with cloze, acceptability rating and self-paced reading experiments on both PP and DP antecedents and sluices in German (see (2) for a sample item). We expect that sluices that correspond to continuations that are more frequent in the cloze study will be more acceptable in the rating study and read faster in self-paced reading. We investigate both effects of the antecedent and of the verb on the likelihood, acceptability and processing of (mis)matching sluices. The likelihood of a second participant for a verb was assessed with a pre-test.

- (2) a. Hans hat mit jemandem getanzt, aber ich weiß nicht, mit wem. PP,MA
Hans has with somebody danced but I know not with whom
b. Hans hat mit jemandem getanzt, aber ich weiß nicht, wer. PP,MM
Hans has with somebody danced but I know not who
c. Jemand hat mit Hans getanzt, aber ich weiß nicht, wer. DP,MM
d. Jemand hat mit Hans getanzt, aber ich weiß nicht, mit wem. DP,MA
e. Hans hat getanzt, aber ich weiß nicht, mit wem. SP,MA
f. Hans hat getanzt, aber ich weiß nicht, wer. SP,MM

Cloze Study We measured the likelihood of (mis)matching sluices with a cloze task. 120 subjects recruited on the crowd-sourcing platform Clickworker completed utterances like (2) ($N = 24$), which were cut off after *nicht*. Responses were categorized as (mis)matches given the antecedent or as unrelated (e.g. *where, why*). The analysis with logistic mixed effects regressions (Bates et al., 2015) shows that matches are more likely (as compared to unrelated continuations) when the second partici-

part is given in the antecedent (sluicing, (2a)) than when it is not (sprouting, (2e)) ($\chi^2 = 52.96, p < .001$). Furthermore, a CONSTRUCTION:VERBBIAS interaction shows that verbs that require a second participant increase the likelihood of a corresponding sluice more strongly under sprouting than under sluicing ($\chi^2 = 104.37, p < .001$). The cloze study confirms our intuitions that (i) explicit antecedents increase the likelihood of a corresponding sluice and that (ii) the verb particularly increases the likelihood of mentioning the second participant when it is not given in the antecedent.

Acceptability Rating Study With a rating task we investigated (i) whether cloze probabilities are reflected in acceptability, and (ii) whether mismatches under sprouting are particularly degraded. We tested both the sluices and the full forms (i.e. including the TP *Hans getanzt hat / mit Hans getanzt hat*) of the materials in (2) in order to tease apart general effects of structural mismatches and effects of ellipsis. 96 subjects recruited on Clickworker participated in the web-based experiment, 48 rated only ellipses and 48 rated only full forms. Items were presented with 60 fillers and rated on a 7-point Likert scale (7 = fully natural). The analysis with CLMMs (Christensen, 2015) in R shows that, among other effects that we discuss in our presentation, (i) mismatches are degraded under ellipsis ($\chi^2 = 43.9, p < .001$), (ii) sluicing is rated as better than sprouting ($\chi^2 = 26.6, p < .001$), and (iii) sprouting improves overall when a second participant is likely given the verb ($\chi^2 = 8.32, p < .01$). As we expected, these effects are in line with the cloze data. Unlike Chung (2006) predicts, mismatches are not particularly degraded under sprouting, as compared to sluicing ($\chi^2 = 0.28, p > .5$).

Self-paced Reading Study We used a self-paced reading study ($n = 48$) to investigate whether cloze probabilities are reflected in processing effort. Subjects read the full forms corresponding to the materials in (2) and 60 fillers in a masked self-paced reading paradigm. We analyzed the residualized log reading times of the *wh*-phrase and the spillover region (*(mit) Hans getanzt hat*), for the *wh*-phrase, with linear mixed effects models (Bates et al., 2015) in R. For both DP and PP sluices, matching sluices were read faster (for DPs, marginally) than mismatching ones ($\chi^2_{2DP} = 3.41, p = .06$, $\chi^2_{PP} = 5.41, p < .05$). DP mismatches under sluicing were read faster than those under sprouting ($\chi^2 = 14.93, p > .001$). Both of these effects are in line with the cloze data, however, there were no effects of the verb bias on reading times.

Discussion Our experiments confirm the central prediction of our processing account: Mismatching sluices are less likely, less acceptable, and harder to process. We find no evidence for a categorial acceptability difference between mismatches under sluicing and sprouting that Chung (2006) predicts. Mismatches are degraded under sprouting, but this results from an the additive effect of a penalty for sprouting and one for mismatches. The cloze data show that, as we expected, the verb also modulates the likelihood of a sluice. This is expected under our processing-based account and questions categorial identity conditions between antecedent and target. In our presentation we discuss possible reasons for why the effect of the verb on sprouting was observed in the cloze and rating, but not in the reading-time data.

Selected references • Chung, S. (2006). Sluicing and the Lexicon: The Point of No Return. Annual Meeting of the Berkeley Linguistics Society, 31(1). • Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In Proceedings of NAACL (Vol. 2). 159–166. • Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3):1126–1177. • Merchant, J. (2013). Voice and Ellipsis. *Linguistic Inquiry*, 44(1):77–108.