

What and when we accept question word coordinations: From creativity to word salad

Sophie Repp (University of Cologne)

May 28

Linguistik Kolloquium, SoSe 2025

English question word coordinations (QWCs) of *wh*-arguments and *wh*-adjuncts like (i) have received very diverse analyses. Strikingly, the diversity also results from different assumptions about which types of QWCs are grammatical. (i) is generally considered grammatical. It contains an optionally transitive verb so either of the *wh*-phrases can be construed with the 'body' of the clause: *what*, *when* do they eat. I call QWCs like (i) flexible due to the flexible argument structure. In non-flexible QWCs, there is either an argument missing (ii), or a superfluous argument (iii) (superfluous subject). Experimental investigations of QWCs with missing objects suggest an influence of the order of the *wh*-phrases: if the *wh*-argument is adjacent to the clause, acceptability is higher. QWCs with missing subjects and with superfluous subjects/objects are judged ungrammatical in the literature but occur in corpora, although superfluous subject QWCs only occur in argument-before-adjunct order (iii). The analyses deal differently with this empirical situation: Some consider the 'ungrammatical' corpus occurrences errors, others provide an account because native speakers produce them.

- (i) What and when did they eat?
- (ii) What and when did they devour?
- (iii) Who or why would you even need this thing?

I present data from a combined corpus-experimental study testing naturally occurring QWCs (corpus) in an acceptability/makes-sense experiment. I extracted all argument-adjunct QWCs from the ententen21 corpus (UK-domains). Of the twelve possible patterns (TYPE (flexible/missing/superfluous) x ARGUMENT (subject/object) x ORDER), ten occurred: there were no adjunct-before-argument QWCs with superfluous arguments, confirming earlier corpus analyses. For the experiment, I selected ten hits per pattern that came with sufficient context plus 40 fillers. 50 UK-English speakers judged whether the QWC (presented with context), sounded natural and made sense (7-point scale). The statistical analysis (CLMM) revealed that flexible QWCs received higher ratings than non-flexible QWCs. Of the non-flexible QWCs, only argument-before-adjunct QWCs with a missing subject showed straight unacceptability. All others showed great variability with medians in the middle or slightly towards the acceptable scale end. Item analysis revealed that non-flexible QWCs improve when the (non)-*wh*-argument is semantically less specified ('who and when individuals entered...' » 'what and how to check the play equipment...'). I propose that non-flexible QWCs are creative production errors which arise through the desire for brevity in a good-enough language production approach. Depending on semantic-pragmatic factors and linear ordering, these errors can be 'repaired' in comprehension, or they are perceived as word salad. Flexible QWCs are analyzed as elliptic biclausal structures.