Focus parallelism precedes lexical parallelism during processing of contrastive ellipsis

Barbara Tomaszewicz & Petra B. Schumacher (Universität zu Köln)

The processing of ellipsis requires accessing the antecedent for the recovery of meaning. One of the current issues in the experimental research on ellipsis is what properties of the antecedent are accessed in memory when the parser encounters ellipsis and how quickly (Phillips & Parker 2014). It has been shown that at the ellipsis site only the relevant information from the previous context, only the antecedent, is reactivated (Shapiro et al. 2003, Snider and Runner 2011) and that the remnant's plausibility given the lexical properties of the antecedent is immediately evaluated (Kaan et al. 2004). It is also known that **lexical parallelism** influences the preferences for the thematic role of the remnant (Carlson 2001, 2002). In (1a-b), in contrast to (1c), the remnant is lexically biased to be interpreted as the subject and object respectively.

(1)   a. Denis saw a movie, not Steven.
        b. Denis saw a movie, not a trailer.
        c. Denis saw Allen, not Steven.

Some types of ellipsis, such as comparatives and ellipsis in coordinate structures, are additionally influenced by **focus parallelism**, signaled by pitch accents or by the presence of a focus particle in silent reading (Frazier & Clifton 1998, Carlson 2003, 2005, 2014, Stolterfoht et al. 2007):

(2)   a. After dinner, only [the JUDge]$_F$ joined the diplomat for coffee, not [the senator]$_F$.
        b. After dinner, the judge joined only [the DIplomat]$_F$ for coffee, not [the senator]$_F$.

Focus resulting from the presence of an embedded question context, e.g. *John wondered who joined the diplomat for coffee*, also affects the congruence of the remnant (Sauermann et al. 2013, Carlson 2014). However, Kertz (2008, 2013) points out that topic/focus parallelism is not specific to ellipsis but rather reflects a general preference for information structural parallelism in coordinate structures and pairs of sentences linked in discourse (Sag 1976, Rooth 1992). It is thus an empirical question how the different parallelism effects, lexical semantic processing vs. higher-level focus structural information, interact during the time course of ellipsis resolution.

Our study: To test the effects of **focus parallelism** we compared **two kinds of context**, an embedded question assigning narrow focus and a wide focus/neutral context.

(3)   a. Sabine wondered who the Englishmen hugged at the party.   ***Embedded Question***
        b. At the party, the Englishmen and the Frenchmen arrived late.   ***Neutral Context***

Our test sentences containing contrastive ellipsis were constructed to counteract the effect of recency, which can be observed in (1c): at the point of ellipsis 'Allen' is the most recently processed information, hence it is more active in memory than 'Denis' and 'Steven' is more likely to be interpreted as an object than a subject. The item's level of activation in memory also results from what has been termed "depth of encoding" (Hofmeister 2007, 2011). For instance, in (2a) 'only the judge' is more semantically complex than 'the judge' in (2b) and thus requires more processing resulting in a deeper encoding in memory. This effect alone could drive the subject bias in (2a) rather than the postulated focus on 'the judge' during silent reading. We chose a modifier which, unlike *only*, can only optionally associate with focus and which does so at a distance, i.e. it cannot assign narrow focus to its syntactic associate. We used the quantifier *meisten* in German, which, as shown in (4a), can be interpreted without any reference to narrow focus (unlike *only*, and unlike the English 'the most Frenchmen'). It can optionally associate with a focus on the adverbial, (4b), or on the subject, (4c).

(4) Die Engländer   haben die **meisten** Franzosen   zur Begrüßung umarmt, …
    The Englishmen   have   the most   Frenchmen for greeting   hugged
    a. [Die Engländer haben die meisten Franzosen zur Begrüßung umarmt]$_F$, …   WIDE FOCUS
      'The Englishmen hugged <u>most of</u> the Frenchmen for greeting.'

b. Die Engländer haben die meisten Franzosen [zur Begrüßung]$_F$ umarmt, … NARROW FOC
'The Englishmen hugged more Frenchmen for greeting <u>than for any other occasion</u>.'

c. [Die Engländer]$_F$ haben die meisten Franzosen zur Begrüßung umarmt, … NARROW FOC
'The Englishmen hugged more Frenchmen for greeting <u>than anybody else did</u>.'

The presence of *meisten* requires a deeper encoding of the object and could thus mitigate the recency advantage of the adverbial, which we tested with **two types of lexical contrast** in the ellipsis remnant:

(5) a. …und nicht [zum Abschied]$_F$. ***Adverbial***   b. ...und nicht [die Spanier]$_F$. ***Noun***
       and not for farewell   and not the Spaniards

Question 1: In the context of the <u>embedded question in (3a)</u> the focus structure is as in (6) and the continuation in <u>(5a) is incongruent</u>.

(6) Die Engländer haben [die meisten Franzosen]$_F$ zur Begrüßung umarmt, …

Sauermann et al. (2013) found that this type of incongruence is detected in early measures in eye-tracking during reading. Will the fact that in our design the adverbial is a potential focus associate of *meisten*, a more recent and lexically parallel contrast facilitate the processing of the (in)congruence?

Question 2: In the <u>wide focus context in (3b)</u>, is the continuation <u>(5b) incongruent</u> due to focus mismatch (in case the focus structure is as in (4a-b)) or due to focus reanalysis (to the subject focus in (5c) which is initially precluded by the context (3b)) or due to an expectation for the contrast with the most recent constituent (so that (5b) violates lexical parallelism)?

Results: We conducted a Self-Paced Reading and an ERP study on the same materials.

Q1: The incongruence with context (3a) and ellipsis (5a) had no effect on the reading times, but elicited an N400 response (negative wave peaking at 400ms after the onset of the adverbial phrase). This indicates a **rapid detection** of a violation of expectations determined by the contextual **focus parallelism** (N400 modulations reflect the scale of accessibility: *given < inferred < new*, Burkhardt 2006, Schumacher & Baumann 2010, Hirotani & Schumacher 2011). This violation, however, appears easy to repair as suggested by a lack of increased reading times and a lack of a P600 effect that indexes discourse reorienting.

Q2: A P600 response (positivity peaking at 600ms) was found with the incongruence in (3b)+(5b), which also elicited higher reading times. The absence of an N400 effect indicates that in the neutral context, there are no strong expectations for the type of the contrast in the remnant, i.e. recency does not drive the incongruence. There is no violation of an expectation for lexical contrast. There is also no reanalysis to subject focus (5c) because such reanalysis requires a reanalysis of the (implicit) prosody of the sentence, which has also been associated with N400 effects (Stolterfoht et al. 2007, Heim & Alter 2006, Cowles et al. 2007, Baumann & Schumacher 2012). The P600 indicates reorienting to include a new referent in the discourse (Stolterfoht at al. 2007, Schumacher 2009, Schumacher & Hung 2012) and the longer RTs show that this process is more costly than with the incongruence in (3a)+(5a). (The reanalysis of the wide focus in (4a) to narrow focus on the lexical contrast 'Frenchmen' may not require a prosodic shift, Féry 2011.)

Our experiments show that contextual **focus parallelism** guides predictive processing and thus affects the **initial stage** of ellipsis resolution (as in Sauermann et al. 2013), while **lexical parallelism** (albeit confounded with recency in (3a)+(5a)) kicks in **immediately afterwards** and facilitates the processing of the mismatching remnant without any extra cost (no P600, no increased RTs). If lexical parallelism had been the dominant/initial factor in the resolution of contrastive ellipsis, there would have been no indication of incongruence either in reading times nor in brain responses. The combined methodology allowed us to track the time course of detection of incongruence with respect to the antecedent's lexical and focus properties.