Testing a functionalist-pragmatic account of constraints on subject pronoun interpretation.

L.Blything[1], A.Theakston[1], S.Brandt[2], & B.Ambridge[1]. [1] = University of Manchester. [2] = Lancaster University.

Background and motivation. A subject position pronoun like '*he*' is almost always allowed to get its meaning from a previously mentioned full NP (e.g., in '***Yusuf*** *said that* ***he*** *is happy', 'he'* could refer to *'Yusuf'*), but is rarely allowed to refer to a full NP that *follows* it (e.g., '***He*** *said that* ***Yusuf*** *is happy'* ['*he*' ≠ '*Yusuf*']). How do English speakers understand when subject pronouns are either allowed or blocked from referring to particular full NPs? This has traditionally been determined by whether or not it violates syntactic rules of language called Principles 'B' (applied to these sentences, a pronoun like 'he' cannot refer to a full NP within the same clause) and 'C' (applied to these sentences, a full-noun-phrase like *Yusuf* must not take its meaning from whatever is in the [main clause] SUBJECT position) (Chomsky, 1981). Impressively, a syntactic account allows relatively rare cases where pronouns could co-refer with a full NP that follows it - specifically the above principles are unviolated when that pronoun appears in a syntactically subordinate clause: '*When he is happy, Yusuf says so'* (*'he'* can refer to *'Yusuf'*). However, syntactic accounts (e.g., Reinhart, 1983:42) have long acknowledged that they can only resolve that '*he*' could mean *either* character in '*Yusuf told Abdul that he is happy*' - here, a functionalist-pragmatic 'add-on' is needed to interpret the ambiguous subject pronoun such that, for example, '*he*' most plausibly means *'Yusuf'* because speakers are more likely to give information about themself being happy than listeners are to receive information about themself being happy (Kehler, 2008). The present three studies advance a functionalist-pragmatic framework (Ambridge et al., 2014; Kuno 1987) by disentangling its predictions from Principle 'C' constraints, reporting that functionalist-pragmatic factors can act not as a mere 'add-on' but instead explain the lion's share for how subject pronouns are interpreted in general.

Method. Under a traditional syntactic approach (Principle 'C'), ***all*** sentences of the form used in our three Experiments (a pronoun like 'he' in the main [first] clause, followed by a full NP like '*Yusuf'* in a subordinate [second] clause) can receive ONLY the interpretation under which he **cannot** refer to '*Yusuf'*, and **must** refer to another character. Such sentence structures were manipulated in terms of their functionalist-pragmatic properties, and rated by native English speakers (*N* = 160 for Exp.1, *N* = 160 for Exp.2, and *N* = 54 for Exp.3) using a 100-point scale for the likelihood that the pronoun referred to '*Definitely NOT Yusuf'* (left), '*Could be Yusuf'* (centre) and '*Definitely Yusuf'* (right). Examples of these manipulations are provided within the results and visualised in Fig.1-3.

Results. The data was analysed using linear mixed effect models. Contra-Principle 'C', participants regularly permitted pronouns to refer to a full NP that follows it even when an initial clause containing the pronoun is a main clause – and even more crucially, this was significantly greater when:

(i) 'Yusuf' *perceived* the second clause event ('*He was waiting in the office when Yusuf* ***[noticed*** *> saw] that the paperwork had vanished')* rather than *communicated* it, since communication verbs – but not perception verbs– imply the presence of '*someone else'* other than '*Yusuf'*, that '*he'* could plausibly refer to (e.g., Exp.1 *b* = 6.76, *SE*= 0.89, *p* < .001, CI [5.02 – 8.50]). For visualisation, see the x-axis of Fig.1-2.

(ii) The real-world event-likelihood of the events described in the first (main) and second (subordinate) clauses suggested that a single person could plausibly be performing both events (Exp.3: *b* = 20.34, SE= 1.37, *p* < .001, CI [17.66 – 23.01]; see x-axis of Fig.3). Consider the example '*He was* ***driving home*** *when Yusuf started* ***indicating***': the alternative interpretation whereby 'he' does not mean '*Yusuf'* may seem unlikely. However, for examples such as 'He *was driving home when Yusuf started cooking'*, it is clear that the alternative interpretation, whereby '*he*' means '*someone other than Yusuf*', is much more likely.

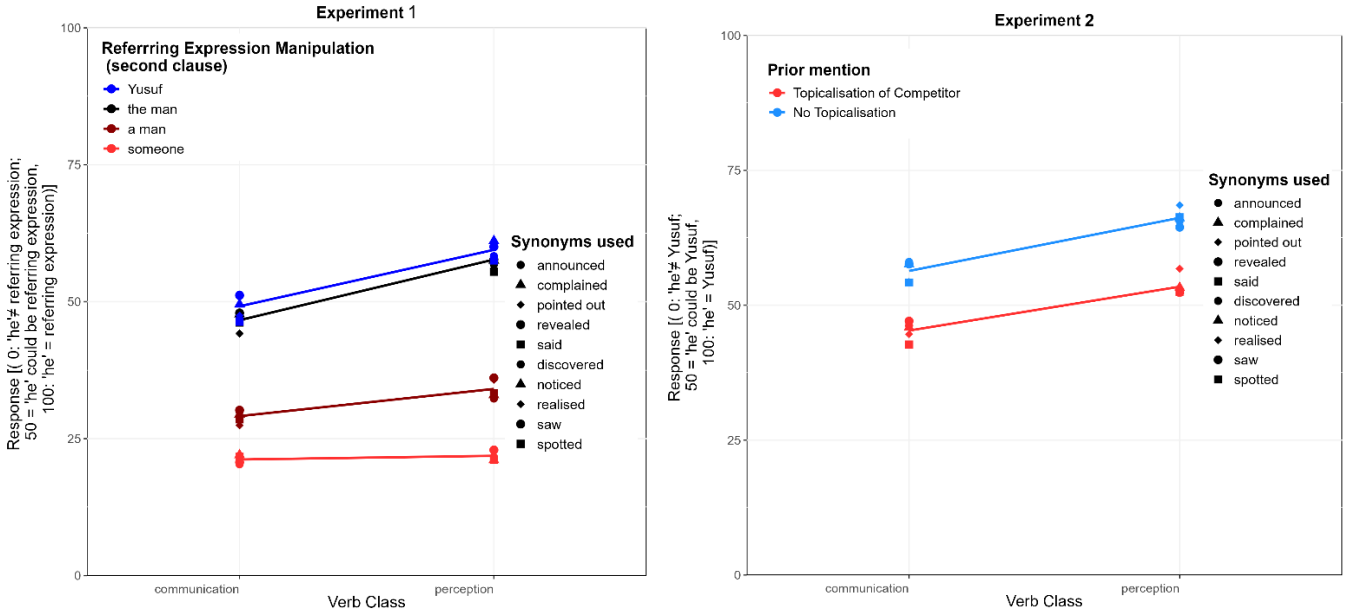
(iii) The full NP of the second clause was of high definiteness ('*He was working on the computer when [****Yusuf****>the man>a man> someone] discovered that the paperwork had vanished'*): Exp.1 *b* = 19.22, *SE*= 1.13, *p* < .01, CI [17.01 – 21.43] (see Fig.1 by-colour conditions).

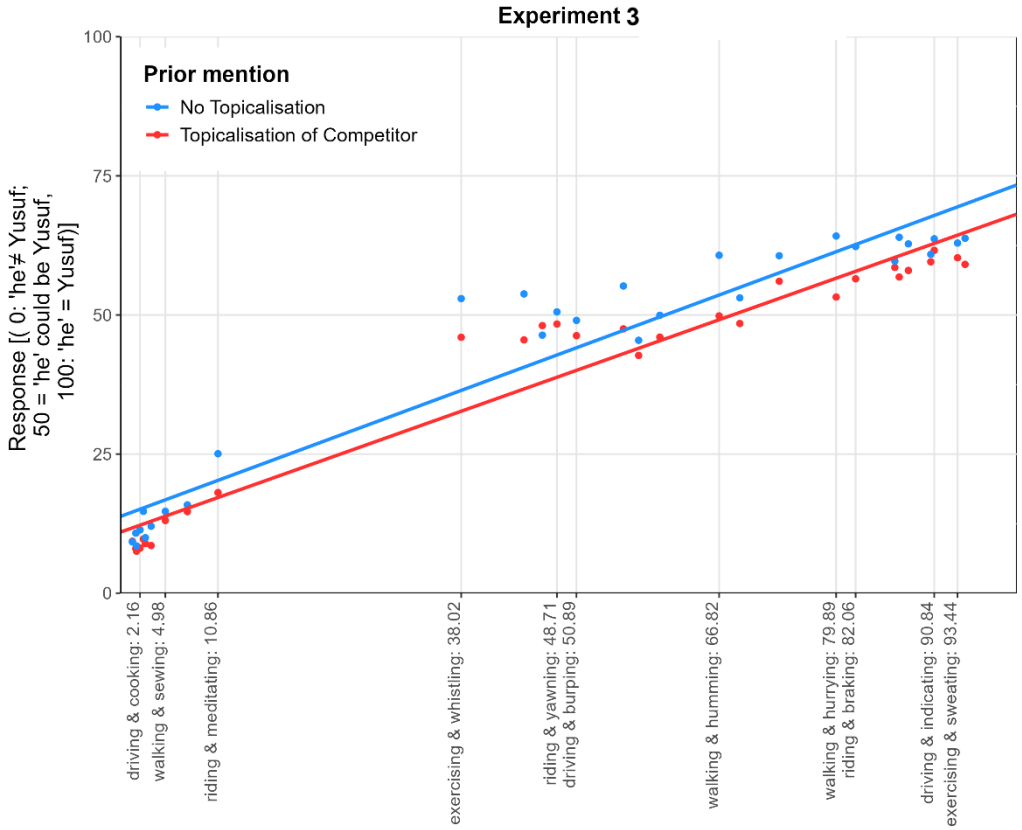
(iv) A prior context mentioned no plausible alternative referent (e.g., *It was Wednesday 2nd August*), rather than naming and topicalizing a plausible alternative referent by first mention (e.g., '*Abdul visited the law firm with Yusuf'*): e.g., Exp.2 *b* = 22.84, *SE*= 1.84, *p* < .001, CI [19.23 – 26.44]; see Fig.2 and Fig.3: by-colour).

Conclusion. Our findings empirically develop the theoretical framework that, once functional-pragmatic factors have been taken into consideration for how subject pronouns like 'he' are interpreted (e.g., which interpretation is a more likely real-world event; who is the topical 'main character' in the ongoing discourse), there may be little remaining for other factors to explain. Together, our findings map onto a single underlying theoretical construct: that 'he' is interpreted as referring to whoever the speaker most plausibly means: either '*Yusuf'* or an alternative character.

Note: ALL Figure y-axis represent the participant response interpretations of *he* [y-axis: scale of 0 to 100 on the extent to which a pronoun refers to the named character of the second clause (100) or not (0)].

Fig.1. Left panel (Exp.1). The effects of verb class event-likelihood (x-axis) and referential-hierarchy (by colour).
Fig.2. Right panel (Exp.2). The effects of verb class event-likelihood (x-axis) and prior-mention (by colour).
Fig.3. Bottom panel (Exp.3). The effects of semantic coherence event-likelihood (x-axis) and prior-mention (by colour).





References. Ambridge (2014). Why universal grammar doesn't help. Chomsky (1981) Lectures on government and binding. Kuno (1987). Functional syntax: Anaphora, discourse, and empathy. Kehler (2008). Coherence and coreference revisited. Reinhart (1983). *Anaphora and semantic interpretation.*