

Null subjects and ellipsis from an idiom perspective

Introduction. Clausal idioms are known to undergo pronominalization and ellipsis in English (Nunberg, Sag & Wasow 1994; Bruening 2015). Here, we report a novel generalization involving ellipsis and a clausal idiom’s subject: in pro-drop and partial pro-drop languages, when a clausal idiom’s predicate is elided, the idiom’s subject *cannot* be realized overtly at all, not even as a pronoun.

We argue that the non-expression of the subject is due to ellipsis, although what gets elided varies across languages. In Chinese, for example, there is a null subject derived via NP ellipsis (cf. Bi & Jenks 2018), while in Brazilian Portuguese, the idiom’s subject is unrealized due to clausal ellipsis.

We suggest that the variation is a *conspiracy*: different ways for a grammar to accommodate the fact that overt pronouns in (partial) pro-drop languages are referential in a way that idiom subjects are not. The data thus also bear on our theories of null subject phenomena and how they might be derived from ellipsis.

Empirical generalization. Clausal idioms can participate in anaphoric relations under specific conditions. In dialogue contexts in English, ellipsis of the idiom VP must co-occur with pronominalization of the subject, and vice versa (1) (see Huang & Mendes 2018). Curiously, in similar contexts in (partial) pro-drop languages, *VP ellipsis must occur without an overt subject*: the idiom subject cannot even surface as an overt pronoun that matches in gender and number; see (2) (Brazilian Portuguese, BP) and (3) (Mandarin Chinese, MC).

- (1) A: Did *the chickens come home to roost*? B: Yes, {**they**/ #the chickens} sure did. [English]
- (2) A: *O pau já tinha comido* quando você chegou lá? B: (#**Ele**/ #O pau/ \emptyset) tinha, mas ...
the stick already had eaten when you got there he the stick had, but
‘Had the stick already eaten when you got there?’ ‘He had, but ...’ (“The stick eats” \approx “A fight breaks out”; note that *o pau* is masculine) [Brazilian Portuguese]
- (3) A: *Wailaide heshang zhende hui nian jing?* B: (#**Tamen**/ #Wailaide heshang/ \emptyset) dangran bu hui!
foreign monk really can chant sutra they foreign monk of.course not can
‘Can foreign monks really chant sutras?’ ‘Of course they can’t!’ (“Foreign monks can chant sutras” \approx “Things from outside are necessarily better.”) [Mandarin Chinese]

Note that this restriction against overt pronoun subjects does not hold in non-idiomatic contexts (4).

- (4) a. A: *O Pedro já tinha comido* quando você chegou lá? B: (**Ele**/ \emptyset) tinha, mas ...
the Peter already had eaten when you got there he had, but ...
‘Had Peter already eaten when you got there?’ ‘He had, but ...’
- b. A: *Xiao heshang zhende hui nian zhe-bu jing?* B: (**Ta**/ \emptyset) dangran bu hui!
young monk really can chant this-CL sutra he of.course NEG can
‘Can the young monk really chant this sutra?’ ‘Of course he can’t!’

Deriving the absence of overt subjects. We assume that the idiomatic reading in (2) and (3) shows that the idiom – including its subject – is syntactically present, only unpronounced. We review two ellipsis hypotheses that generate this outcome: (I) NP ellipsis of the subject, yielding a null pronoun; (II) Clausal ellipsis, which causes the subject to be elided. We argue that both are needed, with data from MC and BP.

I. NP ellipsis of subject, yielding null pronoun. We begin with the English example (1). Following Huang and Mendes (2018), we derive the pronoun subject via NP ellipsis, i.e. *the chickens* \rightarrow *they* (cf. Postal 1969, Elbourne 2001).

We propose extending this analysis to the MC example (3): the acceptable version of (3) contains a null pronoun, derived from NP ellipsis (5). This aligns with independently-motivated proposals by Tomioka (2003) and Bi and Jenks (2018) for null pronouns in East Asian languages. (The VP is elided separately).

- (5) B: --null subj. dangran bu hui! = [_{NP} wailaide heshang] dangran bu hui [_{VP} nian jing]!
of.course NEG can foreign monk of.course NEG can chant sutra
‘Of course they can’t!’ (cf. (3))

The clausal ellipsis analysis does not work as well for MC. Whether MC has clausal ellipsis (sluicing) is controversial (see Adams 2004, Song & Yoshida 2017, etc.). Further, in (3), this analysis entails that the negated auxiliary raises to a position above the subject, which is not attested elsewhere in MC.

II. Clausal ellipsis. The BP example (2) involves affirmation. Building on Martins (1994, 2013) and Kato (2016), we analyze these constructions as verb-stranding clausal ellipsis: the verb raises to a higher focus projection, and the remnant, which contains the subject, is elided.

- (6) B: *Tinha, mas ... = ...* *Tinha*₁ [_{CP} *pau já* *tinha*_T *comido ...*], *mas ...*
had but had the stick already had eaten but
‘He had, but ...’ (cf. (2))

Consider the NP ellipsis analysis where (2) contains a null pronoun subject preceding *tinha*, derived from *o pau* ‘the stick.’ This analysis overgenerates: it predicts that null subjects should be more frequent than they actually are in BP. Matrix subjects positions are often not dropped in BP. Here we see the opposite pattern: the subject *has* to be dropped to maintain the idiomatic interpretation.

Ruling out an alternative: null expletive subjects. For the sake of argument, suppose that clausal idioms enter the derivation as vPs. In (2) and (3), the vP is elided and a null expletive appears as the subject (7).

- (7) B: *Tinha, ... = pro_{exp} tinha* [_{VP} *o pau comido*] ... (cf. (2))

All else being equal, this analysis incorrectly predicts that BP and MC declarative sentences have VSO or Aux-SVO word order. There are also theory-internal concerns about the viability of this analysis: null expletives lack both PF and LF content, and so have no relevance to the interfaces (cf. Sheehan 2006).

Why repeating the idiom’s subject cannot co-occur with VP ellipsis. Following Huang and Mendes (2018), we suggest that repeating an idiom’s subject while eliding its VP leads one to infer that the speaker intends to contrast the subject’s referent with another salient entity. Since no such referent or entity exists in the context of a clausal idiom, the resulting utterance is unacceptable.

Why overt pronoun subjects cannot co-occur with VP ellipsis in BP and MC. We believe that this reduces to some difference between overt pronouns in (partial) pro-drop languages like BP and MC versus non-pro-drop languages like English. For example, using distributional evidence, Bi and Jenks (2018) argue that overt pronouns in MC are referential, while null pronouns are not. We tentatively suggest that overt pronoun subjects cannot appear in (2) and (3) because unbound overt pronouns in pro-drop languages refer to discourse referents. Idiom subjects, however, do not introduce discourse referents in these languages: positively identifying an idiom subject’s referent is often difficult.

Conclusion We presented a novel observation about idioms and ellipsis in (partial) pro-drop languages: idiomatic interpretations are available under ellipsis with the additional condition that the idiom subject cannot be realized overtly. We argued that the absence of the overt subject is derived by ellipsis, although we also argued that there is no uniform ellipsis-based derivation of these phenomenon. In particular, our analysis of the MC data provides novel evidence for Bi and Jenks’/Tomioka’s proposal for null pronouns, and more generally for theories that derive pronouns (null or overt) from full DPs, such as Postal’s or Elbourne’s.

Selected references Bi & Jenks. 2018. Pronouns, radical pro-drop, and ellipsis in Mandarin. *SuB* · Elbourne. 2001. E-type anaphora as NP deletion. *NLS* · Huang & Mendes. 2018. On pronominalization and ellipsis in clausal idioms. *NELS* · Martins. 2013. Emphatic polarity in European Portuguese and beyond. *Lingua* · Nunberg, Sag, & Wasow. 1994. “Idioms.” *Language* · Postal. 1969. “On so-called ‘pronouns’ in English.” In *Modern studies in English*.